DOCUMENT RESUME

ED 048 382 TM 000 502

AUTHOR	Findley, Warren G.; Bryan, Miriam M.
TITLE	Ability Grouping: 1970 -- II. The Impact of Ability Grouping on School Achievement, Affective Development, Ethnic Separation and Socioeconomic Separation.
INSTITUTION	Georgia Univ., Athens. Coll. of Education.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C.
PUB DATE	Dec 70
NOTE	75p.
AVAILABLE FROM	Dr. Morrill M. Hall, Director, Center for Educational Improvement, College of Education, University of Georgia, Athens, Georgia 30601. Identify the title and the part needed (Single copies)
EDRS PRICE	EDRS Price MF-$0.65 HC Not Available from EDRS.
DESCRIPTORS	*Ability Grouping, Academic Achievement, *Affective Behavior, Age Grade Placement, Caucasian Students, *Cultural Isolation, Educational Environment, *Environmental Influences, Flexible Progression, Gifted, Heterogeneous Grouping, Homogeneous Grouping, Longitudinal Studies, Minority Groups, Performance Factors, Research Reviews (Publications), *Socioeconomic Status, Standardized Tests, Student Attitudes
IDENTIFIERS	*Affective Development, Great Britain

ABSTRACT

In this comprehensive review, important studies from the literature relevant to the impact of ability grouping on students are summarized. The following are considered: the effect of heterogeneous and homogeneous grouping practices on academic achievement and affective development; the tendency of ability grouping to produce separation of students along ethnic and socioeconomic dimensions; and the relationship between these dimensions and achievement on tests used to classify children. It is concluded that grouping practices based on standardized tests not only tend to restrict the quality of the instructional experiences of children with respect to academic and social learning, but also, as a result of ethnic and socioeconomic separation, tend to restrict the overall range of experiences and learning opportunities available in the classroom. The report concludes that ability grouping per se is generally ineffectual in improving academic achievement. An extensive bibliography is provided. See TM 000 501, 503, and 504 for other parts of this report. (PR)

ED048382

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL BY MICROFICHE ONLY HAS BEEN GRANTED BY

W. Findley

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

Ability Grouping: 1970

II. The Impact of Ability Grouping on School Achievement, Affective Development, Ethnic Separation and Socioeconomic Separation

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Center for Educational Improvement
University of Georgia
Athens, Georgia 30601

TM 000 502

FOREWORD

In December 1969, a task force was organized for the purpose of advising on the scope and organization of a series of reports regarding ability grouping in the public schools of the United States. Those involved in the planning included:

Warren G. Findley, Principal Investigator

Miriam M. Bryan	Edmund W. Gordon
Paul I. Clifford	Roger T. Lennon
John E. Dobbin	A. John Stauffer
Gordon Foster	Ralph W. Tyler

The Office of Education and the U. S. Department of Health, Education and Welfare were represented by Peter Briggs, Christopher Hagen, and Rosa D. Wiener.

Four documents were planned and have now been completed.

I. Common Practices in the Use of Tests for Grouping Students in Public Schools.

II. The Impact of Ability Grouping of School Achievement, Affective Development, Ethnic Separation, and Socio-economic Separation.

III. Problems and Utilities Involved in the Use of Tests for Grouping Children with Limited Backgrounds, and Alternative Strategies to Such Grouping.

IV. Conclusions and Recommendations

Mrs. Bryan prepared Document I, based on questionnaire responses from schoolmen and supplementary data from Miss Wiener. Dr. Clifford and Mr. Dominick Esposito prepared the basic content of Document II, which was then edited by Mrs. Bryan. Contributions to Document III were secured from Mrs. Bryan, Mr. Dobbin, Dr. Findley, Mrs. Blythe Mitchell, and Dr. Stauffer. The summary and conclusions were prepared by Dr. Findley.

The work presented herein was performed pursuant to a grant from the U. S. Office of Education, Bureau of Elementary and Secondary Education, Department of Health, Education, and Welfare. However, the opinions expressed herein do not necessarily reflect the position or policy of the U. S. Office of Education, and no official endorsement by the U. S. Office of Education should be inferred.

Additional copies of the four documents are available upon request. Write:

Dr. Morrill M. Hall, Director
Center for Educational Improvement
College of Education
University of Georgia
Athens, Georgia 30601

TABLE OF CONTENTS

INTRODUCTION: AN OVERVIEW	1
HOMOGENEOUS AND HETEROGENEOUS GROUPING: DEFINITIONS AND DISCUSSION	4
GROUPING PRACTICES AND SCHOOL ACHIEVEMENT	9
Significant Research Studies from 1960 to the Present	14
GROUPING PRACTICES AND AFFECTIVE DEVELOPMENT	24
Research Prior to 1960	25
Research From 1960 to the Present	26
Ability Grouping - British Style	34
ABILITY GROUPING AND SEPARATION: ETHNIC AND SOCIOECONOMIC	40
The Special Problem of Metropolitan Areas	40
Limited Research on Grouping Practices and Separation	42
Racial Isolation in America	43
Ethnic and Socioeconomic Status in Relation to Test Performance and School Achievement	44
Empirical Consequences of Ability Grouping on Ethnic and Socioeconomic Separation in the Classroom	49
NON-NEGRO MINORITIES	59
SUMMARY AND CONCLUDING REMARKS	61
REFERENCES	64

INTRODUCTION: AN OVERVIEW

The quality of an educational environment may be defined as the quality of the experiences that are provided by that environment. Thus, the extent to which ability grouping tends to enhance or reduce school learning experience is of particular educational significance. If ability grouping tends to restrict the quality of children's school experiences, such practices by design, if not intent, foster an unsound environment for the education of children and should be discontinued. If, on the other hand, evidence suggests that ability grouping tends to maximize the cognitive and social experiences available in a classroom, then such practices should be initiated and/or continued in the interest of maintaining quality education.

Ability grouping is the practice of organizing classroom groups in a graded school to put together children of a given age or grade who have most nearly the same learning achievement or capability, largely on the basis of standardized tests. In the survey conducted as part of the present study, 206 of the 252 school districts reporting the use of ability grouping, or 82 per cent, use standardized tests as an integral feature of the process. (See Table 4 of Document I - Common Practices in the Use of Tests for Grouping Children in Public Schools.) In the discussion that follows all such standardized tests, whether of subject matter achievement, IQ, or "aptitude" will be considered simply different varieties of achievement tests. This terminology is intended to reflect that, functionally, the usual distinction between measures of aptitude and achievement, i.e., innate talents vs. learned talents, is not a meaningful and worthwhile division. In classifying IQ and other aptitude tests, as well as reading, arithmetic, and other subject matter tests as measures of individual achievement, the implication is that a score obtained on each of these instruments reflects the child's level of knowledge in a given subject or skill which, in turn, reflects the environmental and/or developmental end product at a specific point in time.

There are a number of dimensions on which one may evaluate the quality of a particular educational environment. Chief among such dimensions is student achievement in the basic academic skills, i.e., reading and arithmetic. For more than five decades, educators and researchers have focused on these dimensions and have contributed a large body of relevant data.

Recently, a second dimension has received research attention. This dimension can be broadly classified as social learning. Here student attitudes and aspirations, personality development, adjustment to school, social behaviors, etc., are measured to determine in what ways heterogeneous and homogeneous grouping practices influence such affective development.

One would expect that the amount of research in ability grouping would have clarified and settled the basic issues as to which educational practice

contributes more to the academic and social growth of children. Recent surveys suggest that this is not the case.

It is not the purpose of this paper to engage primarily in a detailed review of this research, but rather to present a re-search of literature which documents the finding that grouping practices based on standardized measures of achievement not only tend to restrict the quality of the instructional experiences of children with respect to academic and social learning, but also, as a result of ethnic and socioeconomic separation, tend to restrict the overall range of experiences and learning opportunities available in the classroom.

As will be indicated later, few research efforts have considered the educational relevance of ethnic and socioeconomic variables in the placement of children into ability groups or tracks; few have been directed at the practical consequences of an ability grouping policy on classroom composition with respect to ethnic and socioeconomic separation. Rather, emphasis in the placement of children has resided mainly in IQ, general academic achievement, and reading achievement levels, alone or in combination; with respect to the consequences of ability grouping, emphasis has resided mainly in children's academic achievement, attitudes, and personality development.

Tables 1 and 2 present data which indicate that ability grouping practices tend not to include ethnic and socioeconomic status as criteria in the placement of children, nor are these variables considered in relation to possible de facto segregation along ethnic and socioeconomic dimensions.

Table 1

Number and Per Cent of Studies Using Various Criteria to Determine Ability Group Placement

Criteria	Number of Studies	Per Cent
Academic Achievement	36	72.0
IQ	25	50.0
Reading Level	11	22.0
Teacher Judgment	6	12.0
Sex	5	10.0
Age	4	8.0
Grade Level	3	6.0
Aptitude	2	4.0
Other*	5	10.0
Total	50	XXXX

* Residence, interview, subject marks, interest.

Note: This table apears on page 42 of the NEA Research Summary 1968-S3, Ability Grouping. It is based on 50 selected research investigations of ability grouping.

Table 2

Dependent Variables Used to Test the Effects of Ability Grouping

Dependent Variables Used	Grades 1-6		Grades 7-12	
	Number of Studies	Per Cent	Number of Studies	Per Cent
Academic Achievement	25	93**	21	66
Attitude & Personality Development	9	33	14	44
Social Learning	6	22	10	31
Adjustment to School	4	15	9	28
Teacher Reaction	7	26	7	22
Total	27*	XX	32*	XX

* The totals should not equal the sums of the respective columns since a given study can, and frequently does, appear in more than one dependent variable category.

** Percentages are based on the totals that appear below each column and are rounded to the nearest whole number.

NOTE: This Table was derived from Table 5 of the NEA Research Summary 1968-S3, Ability Grouping.

The balance of this document will be presented in five sections. In the first section, the concepts of homogeneous and heterogeneous grouping will be defined and some findings summarized. The second section will offer a more detailed discussion of relationships between grouping practices and school achievement. In the third section, relationships between grouping practices and affective development will be treated. In the fourth section, since the issue of whether ability grouping tends to separate children along ethnic and socioeconomic dimensions has received considerably less attention in educational literature than its effects on achievement and affective development warrant, evidence bearing on the relationship between ethnic and socioeconomic status and achievement on tests generally used to classify children will be presented. Following this, there will be a detailed discussion of the principal studies which bear directly on the practical consequences of ability grouping on ethnic and socioeconomic separation in schools and classrooms. The final section will be reserved for a summary of the evidence presented in the first four sections and some concluding remarks.

HOMOGENEOUS AND HETEROGENEOUS GROUPING: DEFINITIONS AND DISCUSSION*

In public education, the term "grouping" has been a broad rubric subsuming a wide variety of organizational plans, selection criteria, instructional methodology, and educational philosophies. Since the school has traditionally been defined by its group setting, methods have had to be devised to make the instruction of groups of children more effective and/or more manageable. The major options for vertical organization have been graded, multigraded, or nongraded (continuous progress) schools. Whichever of these plans exists in a school, a concomitant pattern of horizontal organization, which assigns pupils to classes, teachers, rooms, and curricular programs, must emerge.

Homogeneous grouping occurs when classes are formed on the basis of similarity on some specific characteristic of the pupils. The criterion for this classification may be age, sex, social maturity, IQ, achievement, learning style, or a combination of them. The group, however, is homogeneous only with respect to this one criterion, or combination of criteria. In practice, of course, it is impossible to form a group of individuals possessing the identical degree of any characteristic other than sex or other nominal variable like skin pigmentation or eye or hair color, so the objective for homogeneity is that a reduced range of that dimension be represented in the group. Ability grouping is one of the many forms of homogeneous grouping, and generally refers to the use of standardized measures of intelligence, ability, or achievement in a given subject in classifying students into separate ability categories.

When ability grouping is applied to all grades and used throughout a school system, it is usually called "tracking." As applied to secondary schools, children are assigned to clearly labeled curricular tracks, i.e., College Preparatory, Vocational, Commercial, General, or Technical. Practially, this means that for ninth-grade mathematics, a student will be assigned to algebra, business mathematics, or basic mathematics, depending on the track in which he is enrolled. Similarly, students enrolled in the College Preparatory track may be exposed to biology, chemistry, and physics, while Vocational or General students are limited to general science and biology. In addition, students are further channeled into biology for College Preparatory enrollees and biology for General or Vocational enrollees. In short, ability and track-type arrangements tend to divide and separate students for instructional purposes. At the elementary school level, this results in a reduction in the frequency, range, and quality of contacts that a student has open to him; while at the secondary school level, it further means that a student is enrolled in a set program that leads to a set destination or diploma at the end.

If one is concerned with achieving a mixture of children in a given classroom who differ on a number of dimensions, including "ability," a heterogeneous grouping policy can meet this concern. Practically, heterogeneous grouping may be accomplished by either randomly assigning all children

*This section relies heavily on a paper prepared for Dr. Edmund W. Gordon, Director, ERIC Information Retrieval Center on the Disadvantaged, Teachers College, Columbia University, by S. Bernstein and D. Esposito, On Grouping in the Experimental Elementary School Project, November 1969.

in a given grade level or school to the respective classes, such as by choosing alphabetically or choosing every fourth name on a list, or by deliberately structuring classes such that a wide range of ages, abilities achievement levels, socioeconomic backgrounds, and ethnic status, is assured.

It should be emphasized that the homogeneous and heterogeneous grouping concepts are essentially at opposite ends of the same yardstick. Inasmuch as homogeneous grouping can theoretically occur only with respect to nominal variables, it seems evident that homogeneous grouping serves merely to restrict the range of individual differences with respect to certain continuous or ordinal criterion dimensions, while heterogeneous grouping tends to expand the range of individual differences on all dimensions.

The debate between proponents of heterogeneous versus homogeneous grouping has been, in effect, over the issue of ability grouping. Both practices and studies of ability grouping in this country became common in the early 1920's, with the development of standardized group measures of intellectual performance. After a decline from the mid-1930's through the 1940's, there has been a recurrence of interest in ability grouping that has tended to coincide with an increased public concern with academic achievement, particularly in mathematics and science.

The variety of reasons consistently offered with respect to the relative merits of ability grouping are by now well-known to most educators. The rationale for ability grouping, not necessarily based on research findings, generally includes the following points: ability grouping takes individual differences into account by allowing students to advance at their own rate with others of similar ability, and by offering them methods and materials geared to their level; more individual attention from teachers is possible; students are challenged to do their best in their group, or to be promoted to the next level, within a realistic range of competition; it is easier to teach to and provide materials for a narrower range; teachers in heterogeneous groups tend, because of these difficulties, to teach to the average or below-average.

On the other hand, the usual arguments for heterogeneity include these: ability grouping is undemocratic and affects the self-concept of all children adversely by placing a stigma on those in lower groups while giving higher-group children an inflated sense of their own worth; adult life experiences are not ability-grouped, and students must learn to work with a wide range of people; students of lesser ability may profit from learning with those of greater ability; it is impossible to achieve truly homogeneous grouping, even along a single variable, since test data are not generally reliable or valid enough for this type of distinction; homogeneous grouping may provide less sensitivity to individual differences in children by giving the teacher the false sense that students are similar in social needs, achievement, and learning style, while heterogeneity permits different patterns of abilities to emerge within a group of children; and finally, homogeneity tends to segregate along ethnic and socioeconomic lines as well as ability. In commenting on this last

point, Passow (1967) observed that some educators would argue that

> ...ability grouping is simply a means of making respectable the procedures whereby pupils from lower socioeconomic and racial or ethnic minority groups are relegated to the 'slower' and 'nonacademic' programs and provided with a basically inferior education. Observers of racially mixed schools frequently find that ability grouping is a means by which pupils are re-segregated within the school.

And further, Clark (1968) cautioned: "Probably the chief argument against homogeneous grouping is the fact that children so segregated lose their individuality in the education situation.. Homogeneous groupings tend to require that children be seen in terms of group characteristics rather than in terms of their own individual characteristics."

Further arguments and retorts can be put forth for either side of this controversy. One would have hoped that research in ability grouping might have clarified and settled some of these issues--certainly there have been a great many such studies since the 1920's. Several of these studies will be reviewed in the next section of this document. A few summarizing points will be made here.

Firstly, the criteria for grouping students in studies which examine the effects of ability grouping range from various measures of reading achievement to "intelligence," to achievement on the arithmetic concept subtest of the Iowa Tests of Basic Skills. The NEA Research Summary on Ability Grouping (1968-S3) points out a number of the inadequacies of existing measures used as criteria for grouping. Basing groups on IQ tests assumes comparability of mental age and ability, as well as uniform level of abilities, in any one individual. Reading tests may not measure functional reading ability or take into account the variety of factors that influence an individual's reading score. Particularly in young children, it is doubtful that division by ability will be very accurate or valid. Heathers (1969) summarizes the issue succinctly:

> ...students' characteristics as learners are not adequately represented by their scores on a general intelligence test. A student's ease and rate of learning vary greatly from one learning task to another. Also, his level of achievement varies considerably from one curriculum area to another and from topic to topic or task to task within each area.

The dependent measures employed in studies of ability grouping present further problems. Most examine the effects of various grouping practices on academic achievement, measured by standardized tests. Some use measures of attitude and personality development, social learnings, adjustment to school, or teacher reaction. Only a few, however, have used a multivariate approach

to examine differential effects of ability grouping along a number of dimensions. Hence, it has rarely been the case that any of the "common sense" arguments made for or against homogeneous grouping have been tested empirically.

The major purpose of reducing the range of ability in any classroom is, ostensibly, to provide more easily for individual differences. Research studies rarely specify, however, the ways in which instruction is to be adapted or modified from group to group. It is generally implied that either the curricular programs, the methodology, or the pace will be varied. Yet, there appear to be no studies which measure instructional practices, whether these practices are to be kept constant or varied over experimental and control groups.

Goldberg et. al. (1966) summarize some of the many difficulties of interpreting research in ability grouping. They point out that studies vary considerably in their range of objectives, in the basis for determining "homogeneity," in duration, in adequacy of selection bases and means of matching experimental and control groups, in numbers of students, numbers of groups, size of classes, in differentiation of curricula and teaching method, in instruments and techniques used in assessing changes in students, and in the development of teachers for various groups, and have generally failed to examine effects of grouping on teachers and administrators.

If it is assumed that the variables indicated above, either independently or in combination, affect student achievement, then not controlling for these variables in studies of ability grouping tends to minimize the difference in variance between or among ability groups, which tends to reduce the likelihood of finding statistically reliable differences. With this perspective, then, it is not surprising to find that research results are inconclusive. No clear and consistent effects on academic achievement have been found. Effects on pupils' attitudes towards themselves and towards school are also ambiguous. However, regardless of the outcome of any particular study, teacher attitudes invariably favor homogeneous grouping, despite Goldberg's findings that most teachers in their sample were more effective, measured by pupil acheivement, in handling a wide range of ability in only one or two subject areas than in teaching all subjects to one level.

There is some evidence indicating that ability grouping might widen the gap in attainment between rapid and slower learners, gains in higher ability groups being offset by losses in lower ones (Daniels, 1961; Douglas, 1964). Further and more recent studies point to detrimental effects particularly on low-ability groups (Heathers, 1969). Despite the questionable nature of Rosenthal's (1968) data on the effects of teacher expectation on pupil achievement, there is certainly the strong possibility that a "self-fulfilling prophecy" is at work when groups are labeled evaluatively. Conversely, there appears to be little evidence that high-ability pupils suffer in "heterogeneous" classes.

In addition, amidst the wealth of inconclusive outcomes, there appears to be experimental support for what has been a common observation: Heathers

(1969) reports studies which indicate that children of middle and upper socioeconomic status tend to be in the high ability groups, while children of low socioeconomic status and minority groups are in the lower ability groups. A great number of reasons can be offered to explain this consequence. Some of the major causal factors will be examined in later sections. For the moment, however, suffice it to point out that the functional outcome seems to be that ability grouping maintains and increases class stratification in our society (Sexton, 1965).

Eash in a 1961 summary of ability grouping research offers several conclusions that speak to some of the major issues related to homogeneous and heterogeneous grouping practices. These conclusions are:

1. Ability grouping in itself does not produce improved achievement in children.

2. Contrary to statements in previous summaries of the research on the effects of ability grouping on children's achievement..., more recent research evidence seems to indicate that ability grouping actually may be detrimental to children in the average and lower ability groups.

3. Ability grouping at an early age seems to favor unduly the placement of children from the higher socioeconomic class in higher ability groups.

4. Research evidence in the area is quite meager, but what is available does not support the prevalent assumption that college achievement is improved by ability grouping.

5. Ability grouping as an organizational structure may accentuate the attainment of goals, and symbols for goals, of narrow academic achievement to the extent that other broader desirable behavioral goals and objectives are attenuated and jeopardized.

6. The evidence is fairly conclusive that grouping practices in a school can assist in developing social situations that influence the student's perception of self, his sense of dignity and worth, and his attitudes toward other children. In view of this, grouping practices should be concerned with furthering the establishment of social climates that will encourage the intellectual, social and personal development of every child without detrimental effects on individual children.

7. Grouping practices are significant factors in establishing a teaching-learning situation whereby children can acquire the general education skills and abilities needed by all citizens in a democratic society. This means, in brief, that students need opportunities to work in common purpose with a wide range of individuals. Grouping practices which separate students on the basis of ability as

by group IQ or standardized tests reduce th likelihood that students will be exposed to a broader range of ethnic and cultural differences in the society.

8. Pressures to institute certain grouping practices in our schools represent pervasive social problems in our culture. Educators need to be doubly alert that the schools are not utilizing grouping practices which assist in maintaining and promoting social and racial biases which militate against the general education objectives, equal educational opportunity, and the development of each person as an individual.

If the major educational objective of classifying children into restricted range classroom environments is "greater provision for individual differences," and given that there is no clear-cut evidence indicating that this objective has been realized in the tens of thousands of homogeneous classrooms across the nation, then one is compelled to entertain the conclusion that ability grouping, as presently implemented, has failed to establish its merit as a sound educational policy. In this, we second the conclusion put forth in the 1968 NEA report referred to earlier:

> Despite its increasing popularity, there is notable lack of empirical evidence to support the use of ability grouping as an instructional arrangement in the public schools.

The logical implication of these findings is to engineer an educational environment that can practically sustain learning task-oriented small group activities in which more direct individual attention and instruction can be realized.

GROUPING PRACTICES AND SCHOOL ACHIEVEMENT

The literature of better than sixty years relating to research on grouping practices and school achievement has been systematically and thoroughly reviewed by many individuals and groups. Probably the most comprehensive and authoritative reviews have been those of Billett (1932), Ekstrom (1959), Borg (1966), the Research Division of the National Education Association (1968), and three contributors to the Encyclopedia of Educational Research: Otto (1941, 1950), Goodlad (1960), and Heathers (1969).

Each of these reviews is accompanied by an extensive bibliography. Taken together, these bibliographies list hundreds of different studies of the relationship between grouping practices and school achievement. While many of the earlier studies, and some of the later ones as well, would not

be considered today to be truly "research" studies, each of them has information to offer the individual who is interested in pursuing a study of grouping from the beginning.

Billett (1932) reviewed 140 research studies made between 1910 and 1928. He classified 108 of these as "experimental or practical." Of the 108 studies, however, Billett listed only four as "thoroughly controlled"* and two as "partly controlled." Of the four "thoroughly controlled" studies, two were favorable to grouping, one was doubtful, and one was unfavorable. One of the two "partly controlled" studies was favorable to grouping and the other was unfavorable.

Otto (1941) summarized the status of ability grouping as of that date. His conclusions may be summarized as follows: (1) Where adaptations of standards, materials, and methods had been made, the evidence slightly favored ability grouping as contrasted with heterogeneous grouping. However, (2) the evidence of the relative merits of various adaptations of standards, materials, and methods was too inadequate to form a judgment. (3) The greatest relative effectiveness of ability grouping appeared to be for "dull" children, the next greatest for average children, and the least (frequently harmful) for bright children. (4) Evidence regarding particular grade levels or subjects in which ability grouping was especially effective was too inadequate to form a judgment. (5) Most teachers preferred to work with homogeneous rather than with heterogeneous groups. (6) On the whole, parents were favorably disposed to the use of grouping. (7) Although one study showed the great majority of students in schools using ability grouping to be satisfied and happy, evidence regarding the effect of ability grouping on characteristics of students other than knowledges and skills was highly subjective and inconclusive. (8) In general--and this is perhaps Otto's most important conclusion--variability in achievement in ability groups was almost as great (74 to 93 per cent under varying conditions) as it was in unselected groups.

Nine years later Otto (1950) reported that his search of the literature on ability grouping showed no research studies to have been made for 15 years. The conclusions reported by him at this time were, therefore, the same as those reported earlier.

---

*In a controlled study of the effects of ability grouping, the investigator provides evidence that the effects of other possible causes of differences between the groups being compared have been "controlled," i.e., the groups have been matched on the possibly influential variables or statistical procedures have been applied to correct for the possible effects. In an uncontrolled study, the true cause remains in doubt.

Ekstrom (1959) reviewed 33 research studies made between 1923 and 1959. She found 13 studies, with differences having or approaching significance,* which favored homogeneous grouping; 15 studies reporting no differences in achievement between homogeneous and heterogeneous groups, or differences unfavorable to homogeneous grouping; and five studies reporting mixed results, partly favorable and partly unfavorable to homogeneous grouping. Ekstrom could find no consistent pattern for the effectiveness of homogeneous grouping related to age, ability level, curriculum, or method of instruction. She cautioned that the differences in number of favorable or unfavorable studies should not be considered too seriously since the studies differed so widely in quality, purpose, and scope. She noted the inability to control certain relevant factors like the type of teaching and the differentiation of teaching according to ability levels as important weaknesses in most of the studies. She was also critical of the experimental design in several of them, especially the use of matched pairs of subjects based on unwarranted assumptions of similarity in other respects.

Goodlad (1960), who reviewed 12 pieces of literature regarding ability grouping incidental to a review of classroom organization generally, reported conclusions reminiscent of those of Otto 19 years earlier: (1) Evidence with regard to academic achievement appeared to favor ability grouping slightly for slow students and to a greater extent for bright students. (2) The grouping itself was not so significant a contributor to academic achievement as was differentiation by curriculum. (3) Studies of ability grouping in different subject matter areas were somewhat contradictory. (4) Teachers reacted more favorably to teaching homogeneous groups than to teaching heterogeneous groups.

Borg (1966) reviewed 37 research studies made between 1922 and 1962, 20 of them being studies that had also been reviewed by Ekstrom. His

---

*Here and hereafter in this document, the term "significance" will be used in its technical statistical meaning. That is, a difference in favor of one method of grouping or another will be pronounced "significant" if appropriate statistical checks indicate that so large a difference would arise as a matter of chance variation between random samples of the same sizes, so infrequently that it is most reasonable to dismiss this possibility. Instead, it is better to presume that the difference found is attributable to factors that will cause differences in the same direction to occur whenever similar samples are compared that differ in respect to grouping. We speak of differences being "significant at the 5 per cent level" when differences as large or larger would be expected to be found in less than 5 per cent of pairs of random samples of these sizes drawn from a common pool of individuals who had been taught under identical circumstances. We speak with even more confidence of the "significance" of a difference if the likelihood of occurrence of one so large between random samples of these sizes from a common pool is less that 1 in 100; in that case, we speak of the difference as being "significant at the 1 per cent level."

findings confirmed the inconclusiveness found by earlier reviewers to be true of studies on grouping practices and school achievement made prior to the early 1960's. Of the 37 studies, Borg found 20 with differences of significance or approaching significance. Of these 20 studies, 13 were favorable to homogeneous grouping and seven were unfavorable.

The NEA Research Summary, 1968-S3, with 158 bibliographical entries, reports three reviews not covered earlier. Eash (1961) reviewed 28 items. His conclusions have been presented in detail earlier in this document (p. 8). Wilhelms and Westby-Gibson (1961) concluded that (1) there was no evidence that ability grouping per se was leading to improved mastery of subject matter; (2) the evidence slightly favored ability grouping, but the difference was small; (3) if any group had gained from ability grouping, it had been the low group rather than the ablest group; and (4) teachers tended to favor grouping as easing their problems of instruction. Franseth (1964) suggested that the findings reviewed by her raised as many questions as they answered. On the basis of her study, she concluded that factors other than grouping procedures might well account for differences in gains in achievement when they occurred between children homogeneously and heterogeneously grouped.

NEA Research Summary 1968-S3 also abstracted a total of 50 research studies on the effects of ability grouping published since 1960. Twenty-three of the studies were concerned with ability grouping at the elementary school level, that is, in grades 1 through 6; 23 were concerned with ability grouping at the secondary school level, that is, in grades 7 through 12; and four were concerned with ability grouping at both elementary and secondary school levels. Of the 50 abstracts, 42 pertain to the effects of ability grouping on academic achievement. From these 42 abstracts it is possible to infer again that, although the research on the effects of ability grouping on school achievement is extensive, the results, in general, are inconclusive and indefinite; and that factors other than ability grouping account for the differences in achievement that appear when learners grouped according to their abilities are compared with their counterparts in heterogeneous or randomly grouped situations. In this connection, where ability grouping appears to be more successful than heterogeneous grouping, modifications in educational objectives, curricular organization, teaching methodology, and teaching materials may well contribute more to the differences than does ability grouping itself. Some of the research studies abstracted in the NEA Research Summary are described in more detail later in this document.

Heathers (1969), with 84 bibliographical references covering the period from 1932 to 1968 but concentrating particularly on the literature of the 1960's, indicates that the major research studies reported in the 1960's lend strong support to the more recent view that ability grouping is associated with detrimental effects on slow learners, who, when they are placed in low ability groups, have been found to attain lower scores on achievement tests than comparable students obtain when taught in heterogeneous groups. One possible explanation for this phenomenon, Heathers

notes, is that slow learners, in the absence of superior students, have fewer opportunities to learn vicariously through paying attention during classroom discussions in which they can be stimulated by other students. Another possible explanation is the self-fulfilling prophecy, that is, if teachers expect less from students who are assigned to low groups and teach them correspondingly less, the students who are assigned to such groups generally expect less of themselves and behave accordingly; on the other hand, when slow students are assigned to other groups, they are more successful. Heathers also reports evidence that the quality of instruction offered low groups tends to be inferior to that provided groups comprised of abler students. He reports that teachers have indicated that they tend to stress basic skills and factual information with slow learners and use drill with great frequency; conversely, they tend to stress higher levels of conceptual learning with high ability students and encourage them to conduct independent projects. In a national study of the teaching of English in the high school, it was found that teachers tended to employ dull, unimaginative instructional approaches with groups identified as slower learning students.

Heathers also mentions the assumption that ability grouping reduces the range of learning-related differences within a group, and that this reduction of range facilitates teaching and learning. This assumption, however, he explains, tends to be invalidated by the fact that the characteristics of students as learners are not adequately represented by their scores on general intelligence tests. A given student's ease and rate of learning and his level of achievement vary considerably from one curricular area to another, and from topic to topic and from task to task within each area. When students are grouped on the basis of intelligence quotients alone, the range of scores on achievement tests is still great.

Heathers suggests that the most effective way to reduce the range of a class in achievement would be to group differentially subject by subject and to base this grouping on separate measures of achievement for each area. He points out, however, that within such groups there would still remain large differences in ability and many other variables that influence learning.

Heathers also deals with the widely held notion that in ability groups rapid learners are freed from instruction which is geared to less capable students, and that since they are challenged to keep up with their intellectual peers, their achievement is enhanced. Related to this is the further notion that slow learners benefit from instruction geared to their capacities and from experiencing success more often in the absence of abler students. Heathers indicates that these assumptions are of at least questionable validity. He reports evidence that placing a student in a group designated as low or slow stigmatizes the student, and that this is reflected in the student's losing interest in learning and study, thereby further debilitating his achievement.

A direct quotation from Heathers pretty well summarizes the inferences he derives from the evidence he found in the literature on ability grouping he reviewed:

> Writing an epitaph for grouping may well be the task of the reviewer of research on grouping for the 1980 edition of this encyclopedia [i.e., the Encyclopedia of Educational Research]. Even today it appears that grouping as a central theme of organization for instruction has nearly run its course and is in the process of being replaced by a familiar theme--individualized instruction--that became a focus of educational reform in the mid-1960's.

Significant Research Studies From 1960 to the Present

As indicated earlier, NEA Research Summary 1968-S3 contains abstracts of 50 selected research studies on ability grouping which have been published since 1960. Forty-two of these are concerned in whole or in part with the effects of ability grouping on school achievement. The most significant studies will be reviewed again in some detail in this document. In addition, other significant studies not reviewed elsewhere will be reported.

The two most carefully designed and most rigorously controlled studies reported in NEA Research Summary 1968-S3 are those done by Borg (1966) and Goldberg, Passow, and Justman (1966). Both studies were longitudinal, the Borg study being conducted over a period of four years and the Goldberg et al. study for a two-year period.

Borg (1966) used two adjacent and closely comparable school districts in Utah. In one district students were placed in ability groups on the basis of composite scores on an achievement test battery, and an attempt was made to adapt curricular materials to the different ability levels and to adjust the rate of presentation to the level of the individual students. In the other district a program of random grouping with enrichment, that is, an attempt to adjust the depth of learning to individual differences, was employed. In the first year over 2,500 students from grades 4, 6, 7, and 9 were selected for the study; during the second year the sample was increased to about 4,000 students.

In the Borg study, students tested in grade 4 were followed through grade 7; other grade samples were similarly followed over the four-year period of the study. Thus, data were collected from all grades from 4 through 12. The California Achievement Test Battery was used during the pilot study year; Sequential Tests of Educational Progress were used during the final three years of the study.

Borg reported 54 statistical comparisons between randomly grouped and ability-grouped elementary school students. Of the 54, 28 were statistically significant at either the 5 per cent or 1 per cent level; 19 of the significant differences were found to be favorable to ability-grouped students,

while nine favored randomly grouped students. However, since 15 of the 19 significant differences favoring ability-grouped students occurred during the first year of the study, the Hawthorne Effect* apparently operated rather strongly in favor of the ability groups during that year. If the first-year differences had been due primarily to the true superiority of ability grouping over random grouping, the differences would have increased each year as the cumulative effects of the more effective system widened the achievement gap between the two groups. This did not occur; in fact, most of the achievement differences which favored the ability-grouped students disappeared by the time these students had completed the sixth grade.

For elementary school students, Borg reported 18 achievement comparisons where superior students were the foci. Of the 18 comparisons, 11 were statistically significant, with 10 of these 11 favoring ability-grouped students. In terms of overall achievement differences for the four years of the study, ability-grouped superior students were significantly higher than randomly grouped superior students. For average students, however, Borg found no consistent trend favoring either random or ability grouping. In the comparisons between slow students, six significant differences were reported by Borg, with four of the six favoring the randomly grouped slow pupils. When the Hawthorne Effect, which operated on the ability-grouped students during the first year of the Borg study, is taken into consideration, the relatively greater gains of the randomly grouped students are of even greater educational significance. Borg, in this connection, writes: "All in all, we may conclude that neither ability grouping with acceleration, nor random grouping with enrichment, is superior for all ability levels of elementary school pupils. In general, the relative achievement advantages of the two grouping systems were slight, but tended to favor ability grouping for superior pupils and random grouping for slow pupils. As was hypothesized, the differences for average pupils did not consistently favor either grouping treatment."

Since all five of Borg's samples were in junior high school sometime during the four years of his study, it is possible to draw inferences with respect to the relationship between ability grouping in the junior high school and achievement. When the achievement data for the five samples were combined, 60 statistical comparisons between comparable ability-grouped and randomly grouped students were made: 33 in mathematics and 27 in science. Of the mathematics comparisons, five were significant in favor of the ability-grouped students and five in favor of the randomly grouped students, while the other 23 were non-significant. Of the science comparisons, five significantly favored ability grouping and one significantly favored random grouping, while the remaining 21 were non-significant. When Borg's junior high school data were examined for superior, average, and slow ability levels, there was a slight tendency for ability grouping to

*The Hawthorne Effect describes temporary gains that take place because of the novelty of the experimental treatment rather than permanent gains that may take place as a result of the treatment

produce higher mathematics achievement among superior pupils and higher science achievement among average students. Among slow students, random grouping tended to produce higher achievement in both mathematics and science.

Of 30 comparisons made by Borg between achievement in mathematics and science for ability-grouped and randomly grouped pupils in senior high school, only four of the comparisons were significant. All four favored ability grouping, and all four differences were in mathematics achievement: one for superior students, two for average students, and one for slow students. It should be noted that less confidence can be placed in Borg's findings on the high school years than in the elementary and junior high school years because of the relatively small amount of high school data.

From the total Borg data on ability grouping and school achievement, Borg found it possible to state the following conclusions: (1) At the elementary school level, the superior student generally showed greater gains in ability-grouped classes; for average students the pattern of advantages and disadvantages associated with the two grouping treatments was so complex that there was nothing to permit a choice between the two grouping treatments; the slow students generally showed better performance in the heterogeneous classrooms. (2) At the junior high school level, ability grouping led to significantly greater achievement gains for superior students although these differences were not large; for average groups the pattern was somewhat the same, with ability-grouped students making higher achievement scores; slow students in randomly grouped classrooms achieved more than their ability-grouped counterparts. Borg offered these conclusions, however, with the caution that they reflected his own value system and that educators having different orientations might well draw different overall conclusions from the findings of his research. Our conclusion is that his findings may be taken at face value, but with particular note of (1) the large proportion of comparisons (96 of 144) that failed to yield significant differences despite the large samples; (2) the failure of significant differences favoring homogeneous grouping at the end of the first year at the elementary school level to persist or increase therafter; and (3) the fact that whatever modest significant differences favor homogeneous grouping are at the superior level, while low ability level students tend to do somewhat better in heterogeneous groups.

The study by Goldberg, Passow, and Justman (1966) involved about 2,200 students in grades 5 and 6, organized into 15 grouping patterns in 86 classes in 45 New York City elementary schools. The grouping criterion was intelligence, and five ability levels were designated: (a) gifted, IQ 130 and over; (b) very bright, IQ 120-129; (c) bright, IQ 110-119; (d) average, 100-109; and (e) low or below average, IQ 99 and lower.

The authors set out to investigate three null hypotheses: (a) The presence or absence of extreme ability levels (gifted or slow) has no effect on the changes in performance of other ability levels. (b) Narrowing the

ability range in the classroom has no effect on changes in the performance of students. (c) The relative position of any ability level within the range has no effect on changes in the performance of students. The hypotheses were tested for five major variables: (a) academic achievement, (b) self-concept, (c) interest and attitudes toward school, (d) assessment of more and less able peers, and (e) teacher ratings of students. Only the first of the variables will be discussed here; the others will be discussed later in this document.

In general, the results showed that in predominantly middle-class elementary schools, narrowing the ability range in the classroom on the basis of some measure of general academic aptitude will by itself produce little positive effect on the academic achievement of students of any ability level. In contrast, presence of gifted students in a class tends to raise science achievement of all levels of pupils, while presence of low ability students has a similar positive effect on arithmetic achievement.

Assessment of the various ranges of grouping patterns showed the broadest pattern to be generally somewhat more effective than any of the combinations of patterns with narrower ranges. A most significant finding was that gains in achievement were more strongly influenced by teacher differences and group differences in individual classrooms than by the presence or absence of high ability students, the range of ability in the class, or the intellectual ability of the students. Between-class variability was greatest for the gifted students and least for the slowest students. When teacher effectiveness across ability levels was analyzed, it was found that teachers were more effective in teaching one or two subjects to a wide-range ability group than in teaching several subjects to a narrow-range ability group. In fact, most teachers were more effective in teaching one subject to several ability groups simultaneously than in teaching all subjects even in narrow-range classes. Finally, average achievement across all subjects was greatest in classes including four or all five of the ability levels originally described in this summary.

Other longitudinal studies on ability grouping have been reported by DeGrow (1963), Kline (1963), Morgenstern (1963), and Tobin (1965).

DeGrow (1963) conducted a study in Port Huron, Michigan involving a three-part research design. The criterion was reading achievement as measured by the California Achievement Test. In a one-year study, two groups of students in grades 4, 5, and 6, matched on the basis of IQ, grade level, sex, and reading scores, were involved. One group was taught in a homogeneous setting, with vertical grouping* according to reading level; the other, in heterogeneous classes. At the end of the year, there were no dignificant differences in achievement between the homogeneous and the heterogeneous groups,

*For vertical grouping, students in grades 4, 5, and 6 were assigned to reading classes on the basis of reading level rather than by grade.

even though variation in reading grade equivalents had been reduced from 8.0 to an average of 1.13 through the homogeneous grouping. In a four-year cross-sectional comparison, comparative data collected for two preceding years indicated that vertical grouping did not make a difference in the average reading achievement gains of students. In a three-year longitudinal comparison, mean reading gains for 180 students who had remained in the homogeneous groups through grades 4, 5, and 6 were not related to this method of grouping. It was DeGrow's conclusion that vertical ability grouping in reading in grades 4 through 6 did not contribute to gains in reading achievement.

Kline (1963) evaluated the tracking plan in St. Louis public high schools. An experimental group was tracked over three to four years, while a control group was traced through their results in heterogeneous classes over the same period. The two groups were matched initially on the Iowa Tests of Basic Skills. The final criteria were teachers' marks and scores on the Iowa Tests of Educational Development. On teachers' marks, 40 experimental-control comparisons were made. For four of the 40 comparisons, the experimental group was higher; for five of the comparisons, the control group was higher. On the tests there were 36 experimental-control comparisons. For four of these, the experimental group was significantly higher; for seven, the control group was significantly higher. Kline concluded that tracking appeared not to make much difference in the achievement of St. Louis public high school students.

A group of sixth graders who had been in homogeneous (ability-grouped) classes for a three-year period were compared with a group that had been instructed in heterogeneous classes over the same length of time, by Morgenstern (1963). The measures used were the Stanford Achievement Test, the California Test of Mental Maturity, the California Test of Personality and Thinking About Yourself. Morgenstern's major conclusion was that ability grouping does not result in significantly greater increments in overall academic achievement or in significantly better personal and social adjustment than does heterogeneous grouping. One of her important subfindings was that in certain specific subject areas, such as language and word meaning, the homogeneous group was significantly superior. For the lowest IQ groups, those grouped homogeneously showed the greater gain in academic achievement.

Tobin (1965) reported a study involving students from grades 1 through 6. The study, covering an eight-year period, included a heterogeneous control year, 1954; a transition year, 1955; and six experimental years, 1956-1961. During the experimental years, students were grouped yearly within each grade on the basis of reading ability; similarly, each year the grades were divided into thirds on the basis of IQ. Each experimental year was compared with the control year. Tobin found that the total group, each of the three IQ level groups, and every separate grade maintained stability in mean intelligence over the eight years. The total group showed positive upward trends in reading and in general achievement; the same was true for the high, average, and low ability students. There was an upward trend in general achievement that was significant in all grades except the third. For reading, gains

were significant in grades 1, 2, and 6. Tobin believed there was no Hawthorne Effect in his study, inasmuch as the greatest increases took place in the later years of the study.

A number of single-year studies of ability grouping, some involving single grades and/or single subjects and others involving several grades and several aspects of student achievement have been reported in the past ten years. Studies by Provus (1960), Fick (1962), Loomer (1962), Mikkelson (1962), Matzen (1965), Flowers (1966), Peterson (1966), Wilson (1967), and Maynor (1970) have been selected for review here.

Provus (1960) studied 494 students in grades 4 through 6 in Homewood, Illinois. Homogeneous classes made up of students grouped for arithmetic only--the academically talented, average students, and slow learners--were compared with heterogeneous classes. On the basis of results on the arithmetic concepts subtest of the Iowa Tests of Basic Skills Provus concluded that children at all ability levels, grouped by ability, were more familiar with arithmetic concepts and fundamentals than children who were not grouped according to ability. He further concluded that the academically talented students profited most from ability grouping; the average students profited slightly; and the slow students profited no more from homogeneous grouping than they did from heterogeneous grouping.

Grade 7 students in Olathe, Kansas, were studied by Fick (1962). He formed homogeneous and heterogeneous classes which were pre-tested and post-tested with the Iowa Tests of Basic Skills, the Index of Adjustment and Values, General and Test Anxiety Scales for Children, and the Scale of Attitudes toward the School Situation. Fick found that his students in homogeneous groups averaged no differently on achievement tests than those in heterogeneous groups. The low ability students in heterogeneous classes were superior to those in homogeneous classes in reading comprehension and punctuation. High ability students in homogeneous classes scored higher on uses of references than did those of similar abilities who were taught in heterogeneous classes. The homogeneous-heterogeneous comparisons on the other instruments will be discussed later in this document.

Loomer (1962) conducted a study involving 490 students in grades 4, 5, and 6, enrolled in 23 different classes. Five heterogeneous classes contained all levels of ability. The homogeneous groups included a high group and a low group. The homogeneous high group contained all ability levels except low students; the homogeneous low group contained all ability levels except bright pupils. The achievement growth from February of one year to February of the next year was measured by the Iowa Tests of Basic Skills. Loomer reported no significant differences between homogeneous high and heterogeneous groups except for vocabulary at grade 5 in which the homogeneous high group was superior. No significant differences were found between homogeneous low and heterogeneous groups. No significant differences between homogeneous high and homogeneous low groups were found except in grade 4, in language and total achievement, and in grade 5, in vocabulary and total achievement, where the homogeneous high

arrangement produced superior results. No significant differences were found on any test between homogeneous and heterogeneous classes insofar as bright level students were concerned; for the low ability students, the only significant differences in achievement were found in grade 5 in reading and in grade 6 in language, where the heterogeneous grouping proved superior. Loomer concluded that his evidence indicated no decided advantage to homogeneous grouping over a random method of assigning pupils to classes.

Mikkelson (1962) studied 280 students of superior mathematical ability in grades 7 and 8 in a Minneapolis junior high school. One hundred forty of the students were studied during the 1958-59 academic year; the other 140, during 1959-60. Thirty-five students in each grade, assigned to one homogeneous class, on the basis of mathematics achievement, Otis IQ, and teacher judgment, comprised the experimental group; the control group was comprised of 35 students placed in traditional heterogeneous classes in each grade. During the first year, no special adjustment in curriculum was made; in the second year, the curriculum was adapted to the homogeneous group by means of acceleration. Mikkelson reported that no differences in mathematics achievement resulted from grouping students of superior mathematical ability when no adjustments were made in the teaching procedures or the curriculum; but that with an accelerated curriculum, the homogeneous group accomplished more than those regularly grouped.

Matzen (1965) studied 1,100 black and white students in grades 5 and 7 in 11 different schools in the San Francisco Bay area to determine the relationship of the proportion of black children in a classroom to the mean scholastic achievement of black and white students. Test findings showed a tendency for both achievement and IQ to vary inversely with per cent of black students, with, however, numerous exceptions; achievement varied directly with socioeconomic level; when IQ and socioeconomic status were held constant, achievement tended to fall as the per cent of black students rose, but the tendency was not strong enough to reach statistical significance. In the fifth grade, where students were less homogeneously grouped than in the seventh grade, the black-white differentials in achievement were greater. In the seventh grade, with bright black children and bright white children in the same classrooms, black-white differences in achievement were minimized. Matzen's findings may be interpreted in many ways, but it is perhaps best to note that they are consonant with those of McPartland's more substantial study discussed later in this section.

Flowers (1966) tested what is commonly called the "self-fulfilling prophecy."* He hypothesized, "If one of two groups of students of similar

*Heathers (1969) cites the study by Rosenthal and Jacobson (1968) as the most dramatic evidence of the self-fulfilling prophecy. In that study, randomly selected students from a class were identified to the teacher as "academic spurters." Over the next several months, these students showed reliable gains in IQ scores, a finding that was equally true of students who were in fast, medium, or slow groups. Unfortunately for this viewpoint, that (continued, p. 21)

tested ability and achievement is assigned arbitrarily to a moderately higher level section and is taught on that level for a year, the group so placed will surpass the other group in tested achievement by the end of the academic year." Flowers worked with seventh-grade students in two experimental groups and two control groups matched on scores on achievement and intelligence tests. The two experimental groups were shifted to higher section designations than their test data would have warranted without their knowledge or their teachers'. Despite a slight trend to higher achievement for the experimental groups, Flowers concluded that his hypothesis was not validated. Extraneous uncontrollable factors evidently operated in this research, such as community differences, school assignments, and teacher styles. It appeared possible to Flowers that the upward trend was related to teacher expectation since a questionnaire indicated that teachers of the experimental groups favored the "high" ability groups, were more sensitive to the need for remedial instruction, and made greater attempts to motivate the "high" ability groups.

Peterson (1966) studied students in grades 7 and 8 in a junior high school in Chisholm, Minnesota. These students were grouped in three ability levels--high, middle, and low--on the basis of six tests of scholastic ability. One half of the students at each level were taught in homogeneous groups; the other half were placed in matched heterogeneous sections. Eight achievement tests were given at the beginning and the end of the year in order to measure growth. At the end of the year Peterson studied differences in the groups in achievement and attitudes toward school. All comparisons that showed significant differences between the groups--and the majority of these were for arithmetic achievement--favored the heterogeneous groups; but only three of the 24 comparisons at grade 7, and eight of the 27 comparisons in grade 8, were statistically significant. Peterson concluded that his study "failed to offer sufficient support for the superiority of either homogeneous or heterogeneous grouping."

Wilson (1967), in a study in Richmond, California, found that large initial differences in the "social inheritance" of children entering school are not perceptibly ameliorated by standard school programs of remedial reading, special classes for the mentally retarded in segregated schools, and grouped classes within schools. Socially isolated schools failed to remedy the learning deficits of lower class children, both black and white. Other findings of Wilson's extensive research will be reported later in this document.

Maynor (1970) compared achievement of 680 black, 127 Indian and 608 white students before and after the first year of integration in grades 6 through 12 in Hoke County, North Carolina. The slopes of the regression lines for achievement, relative to grade placement, on the reading, language, mathematics and total scores on the California Achievement Test showed no change, so it was possible to compare differences in average achievement over the range of grades. Blacks showed gains in all parts, but only those in mathematics and total scores were significant at the 5 per cent level. Indians and whites showed neither gains nor losses. Blacks did their best when taught by Indian teachers.

(con't. from p. 20) study and further argument by Rosenthal (1970) involve questionable statistics (Thorndike 1968, 1970) and several efforts at replication have proved unsuccessful (Barber, et al., 1969).

Two significant analyses of data from the report Equality of Educational Opportunity, principally authored by Coleman (1966), are extremely pertinent to any current discussion on ability grouping and school achievement. The first of these is the work of McPartland (1969), a colleague of Coleman's on the Educational Opportunities Survey, which resulted in the most comprehensive body of data ever collected on public schools and their students in the United States. The second is by Mayeske (1970), charged with colleagues at the U.S. Office of Education with the responsibility of illustrating and documenting the structure and functioning of the American public school system.

McPartland (1969) analyzed data on students from a sample of schools selected from metropolitan areas of the New England and Middle Atlantic states participating in the Survey. He studied 5,075 ninth-grade black students who had attended their present schools in the previous years, using three variables to set up cross classifications: a six-level family background scale constructed from students' reports of their mothers' education and students' responses on a nine-item check list of possessions in the home; the per cent of white students in the ninth grade of a student's school, partitioned into four categories; and four groupings derived from the student's report of the proportion of his classmates who were white. Average achievement scores on a 60-item test of verbal ability derived from the School and College Ability Test were calculated within cells of cross-classification of the variables used. Summary measures were then derived from McPartland's cross tabulations. From the analysis of ninth-grade students in the metropolitan Northeast, McPartland concluded that the potential favorable effects of school desegregation on black achievement can be offset by segregation within the school. He found that only black students in mostly white classes demonstrate any added achievement growth due to attendance at mostly white schools. On the other hand, he found, class desegregation has a favorable effect on black student verbal achievement, no matter what the racial enrollment of the school. He provides evidence that the differences in verbal achievement between black students in mostly white classes, and black students in mostly black classes, cannot be explained by selection processes which operate within a given school. The profound effects of ability grouping on ethnic and socioeconomic separation in the public schools will be described later in this document.

Mayeske (1970), in an article presenting excerpts from a report entitled Our Nation's Schools, indicated that very little of the influence of the schools on achievement can be separated from the social background of their students and very little of the students' social background can be separated from their schools. To quote Mayeske in his discussion of the data at hand:

> Although the relationships are not large, those aspects of the schools involved in student achievement pertained to the teaching staff's verbal skills, racial-ethnic composition, salary level, special staff and services, and their view of their teaching conditions. . . . When the backgrounds of individual teachers were examined, the existence of a dominant color-caste system in the preparation of teachers was discovered. The self-perpetuating

> role that it could play through the reinforcement of differential verbal skills along racial-ethnic lines was suggested whereby teachers tend to teach students from the same socioeconomic and racial-ethnic background as their own.

Mayeske also indicated that students from the higher socioeconomic levels, who have an intact family structure and happen to be white or oriental, enter school with more fully developed skills and motivation which enable them to benefit more from their educational opportunities than their less privileged counterparts.

How ability grouping may serve to widen the achievement gap between black and white students as a result of further separation in schools and classrooms resulting from the grouping will be discussed later in this document.

It is interesting to note that while the great debate has been going on in the United States during the 1960's over the relative merits and demerits of ability grouping or "tracking," a similar debate has been taking place in England over their ability grouping or "streaming" system. Since, however, most of the significant research that has been done in England has been concerned with the effects of "streaming" on the social and personal development of children rather than on their academic attainments, the pro's and con's of streaming, as the English see them, will be discussed later in this document.

A brief summary note regarding the effects of ability grouping on school achievement is that (1) separation into ability groups, when all children involved are considered, has no clear-cut positive or negative effect on average scholastic achievement, and (2) the slight trend toward improving the average achievement of high level groups is offset by a substantial loss by average and low groups. How these effects may be produced by the fact of ethnic and socioeconomic separation resulting from ability grouping is the subject of a later section of this document.

One special footnote is a trend in the results of ability grouping nowadays as contrasted with findings in the 1920's and 1930's. The earlier studies more often than not reported gains by the low groups and losses by the high groups when compared with similar students taught in heterogeneous classes. Today, the trends are just the opposite: any advantages are shown by high level groups, disadvantages are shown quite commonly for the low groups. Why?

A possible explanation is that in the earlier period, strong academic motivation was accepted as a favorable characteristic of individuals, to be prized when noted, but not to be expected under the prevailing drill emphasis in instruction, while the current concept of a "dropout" as one deprived unfairly was yet to be born; currently, since Sputnik in 1957, strong academic motivation and achievement have been "demanded" by our

technological society, especially through middle-class parents, with concomitant wide acceptance that lack of this composite of achievement and motivation in minority groups is a fundamental source of deprivation. The "low" feel low and behave ineffectively to secure the benefits in upward mobility that education provides.* All of which leads naturally to the discussion of the impact of ability grouping on and through affective development.

GROUPING PRACTICES AND AFFECTIVE DEVELOPMENT

Ability grouping has probably been criticized more frequently with respect to its emotional and social aspects than with respect to its effects on academic achievement. Many opinions have been hazarded concerning emotional and social results, but the research evidence, at least until very recently, has been thin indeed, perhaps because emotional and social growth are more difficult to assess than intellectual growth.

Of the 33 studies reviewed by Ekstrom (1959), only one touched upon the social and personal adjustment of homogeneously grouped students. Byers (1961), reviewing the literature from 1930 to 1960, found only eight studies having to do with emotional and social growth, made prior to 1960, that were worthy of review. Borg (1964) included among his references eight studies made prior to 1960 that were concerned with non-cognitive variables; most of these were the same studies reported earlier by Byers. Of the 50 abstracts of research studies made since 1960, presented in the NEA Research Summary (1968-S3), 15 are concerned, in whole or in part, with social and personal adjustment. The contributors to the Encyclopedia of Educational Research--Otto (1941, 1950), Goodlad (1960), and Heathers (1969)--have had little to report on the relationship between grouping practices and affective development. Even Heathers lists fewer than a half dozen research studies concerned with this aspect of grouping.

As there has been little uniformity of opinion regarding the effect of ability grouping on the social development of students, just so has there been little uniformity among the findings reported for the research studies that have been made. However, while the literature concerning the social aspects of ability grouping includes at least some evidence to support any stand one might take, much of the evidence, especially the more recent evidence, seems not to support the generalization that grouping

*Today, when "all the children of all the people" are in school up to a compulsory attendance age limit, the low achieving groups contain far more children of minority and low socioeconomic groups than earlier, when the comparisons were between groups within a narrower range of socioeconomic and ethnic variation.

students according to ability contributes to the development of desirable attitudes and healthy self-concepts, especially among slow learners.

A number of the most significant research studies concerned with grouping practices and various non-cognitive variables--self-concept, attitudes, interests, sociometric patterns, personality traits--deserve to be noted here. Some of these have been reported by previous reviewers of the literature; a larger number have not. Because several of the studies were concerned with more than one variable, the studies are reported in chronological order rather than by aspect of affective development.

Research Prior to 1960

Luchins and Luchins (1948) interviewed 190 children in grades 4, 5, and 6 of a New York City public elementary school. They found that a high percentage of the students in the bright, average, and dull classes preferred to be and believed their parents would prefer them to be in the higher section of their grade rather than in the lower section. While most of those who were in the bright classes indicated that they would be unwilling to give up their higher class status even if the teacher of the lower class were "better and kinder," a majority of those in the dull and average classes would have been willing to change their class because of the teacher factor. A high percentage of the children in the bright group did not frequently play with nor would they choose their best friend from among students in the less able class; while most of those in the average and dull groups were willing to choose playmates from the brighter group and showed a willingness to select best friends without regard to the identification of their class. Many dull students felt inferior and ostracized, and believed that there was stigma attached to the dull class level. There was strong social pressure to be in the higher class. The brighter children, in turn, were, on the whole, snobbish in their attitude toward those who were in the lower class. The Luchins concluded that homogeneous grouping seemed to help create a kind of caste system in the school.

Justman (1953) compared two groups of gifted high school students in New York City, matched on the basis of school attended, grade, sex, mental age, chronological age, IQ, and achievement in reading and computational skills. The experimental groups were special rapid progress classes; the control groups were in heterogeneous normal progress classes. On the basis of results on a variety of tests, Justman concluded that segregation of gifted children in special progress classes is accompanied by academic achievement superior to that attained by matched students in normal progress classes with no detriment to social acceptance, interests, attitudes, and aspects of personality.

Horace Mann (1953) studied gifted children in grades 4, 5, and 6 in Pittsburgh. These children spent half of the school day with typical children in art, music, and physical education classes; the other half of the day was spent with other gifted children in classes devoted to academic learning and enrichment programs. Mann sought to determine how real were the friendships between gifted and typical children in this program of partial segregation; he also attempted to measure the social position of gifted children among their gifted classmates. He found

that the gifted children chose as friends other gifted children more often than they chose typical children; typical children preferred other typical children as their friends. Rejections followed the same pattern. Mann concluded that grouping heterogeneously for part of the day did not produce the desired mingling among children of various ability levels. Acceptance and rejection were stronger within an ability group than between groups.

Luttrell (1958) studied 27 sixth-grade students in Greensboro, North Carolina with IQ's of 130 or above in a special class (experimental) and a comparable group scattered among eight classrooms (control). Both the experimental and control groups were tested in the fall and the spring with an achievement test, the Mental Health Analysis, and the Social Traits Rating Scale. The results on the Mental Health Analysis Scale showed no difference between the two groups, both groups making a slight gain during the year. On the part of the Social Traits Rating Scale based on teacher ratings, the groups were highly similar in November, but by May the control group showed greater incidence of these undesirable traits: boastful, bossy, noisy, sulky, quarrelsome. The part of the scale filled out by the students revealed a high degree of acceptance of the gifted child in the regular classroom. While the number of students was small and the time involved in the study short, the results generally favored the homogeneous group.

Goldworth (1959) studied a program in which gifted children in grades 4 through 8 in a suburban community in the San Francisco Bay area were assigned to special grouping for three hours a week. The 63 classrooms containing fast learners were randomly divided by school and by grade level into experimental and control groups which were comparable in size, IQ distribution, number of learners, and "degree of acceptance." Pretests and posttests, including the Columbia Social Distance Scale and three sociometric tests, were administered to all students. Goldworth found that the program had a limiting effect on the number of classmates whom children accepted as best friends, but had no effect on fast learners' acceptance of classmates as best friends, on group cohesion, or on subgroup preferences. The proportion of children who showed an increase in the degree to which they were accepted as friends by their classmates was significantly greater in the control groups. While this study is widely referred to in the literature, the results should be interpreted with caution since they were based on a study of somewhat less than five months in duration.

Research from 1960 to the Present

"Is ability grouping good in the way children look at themselves?" "Is it good in the way teachers look at children?" Maxine Mann (1960) studied 102 fifth-grade children through the use of self-reports. The children had been classified into four ability groups upon entering first grade on the basis of results on group intelligence tests and reading readiness tests, but were officially labeled only by teachers' names. Two of the questions children were asked to answer were pertinent to the study: "Which fifth grade are you in?" "How do you happen to be in this particular fifth-grade group rather than some other?" Mann found that the highest and lowest groups were most aware of the

level of grouping, identifying their groups as "high fifth," "high," "best," "top fifth," and as "low fifth," "low," "lower" rather than by the teacher's name. The reasons the children gave for their assignment to their particular groups helped to bring their self-pictures into clearer focus. "I'm smart," "We're smarter," "I'm too dumb," and "We don't know very much," "We are lazy" account for more than half the answers to the second question. In the top section, all the children gave positive responses in terms of ability or achievement and no negative responses. In the second section, all the responses were still positive although only about one-fourth of them were in terms of ability or achievement. Most children in the third section and all of the children in the lowest section gave responses that indicated negative, or unfavorable self-concepts. Mann's deduction was that ability grouping is cruel to all but the top students.

In a study of gifted children in California, Simpson and Martinson (1961) administered the California Psychological Inventory to 115 students in special class groups and 56 comparable students given classroom enrichment or acceleration at the eighth-grade and high school levels. The special classes made significant gains in 19 instances and significant losses in three instances on the Inventory, while the other students made significant gains in nine instances and significant losses in eight. Eighth-grade boys in the special classes made significantly greater gains than the other boys in Self-Acceptance; eighth-grade girls in the special classes made significantly greater gains than the other girls on Self-Acceptance and Flexibility; high-school boys in the special classes made significantly greater gains than did the other boys on Social Presence and Tolerance; and high-school girls in the other groups made significantly greater gains than the special class groups in Social Presence.

Fick (1962), in his study of seventh-grade students in Olathe, Kansas, previously cited, used the Index of Adjustment and Values, the General and the Test Anxiety Scale for Children, and the Scale of Attitudes toward Social Situations, along with an achievement battery. Classes grouped homogeneously and heterogeneously were pretested and posttested with all four instruments. As with achievement, the homogeneous and heterogeneous comparisons showed no significant difference in changes in peer behavior, learning needs, teacher-pupil relationships, or self-concept. Responses of students to the anxiety scales, however, indicated significant increases in both general and test anxiety on the part of the ability-grouped students.

Locke (1962) studied the effect of separating rapid learners (Otis IQ and achievement above 89th percentile and a grade-point average above 3.8) from non-rapid learners. One group of rapid learners were homogeneously grouped in one class and all other students were heterogeneously grouped. In the control group, all students were heterogeneously grouped. Seventy-five matched pairs of rapid learners and 193 pairs of non-rapid learners were studied over a two-year period. The Iowa Tests of Basic Skills and the California Test of Personality were given at the beginning and at the end of the interval. While both rapid and non-rapid learners homogeneously grouped made significantly

greater achievement gains than did those heterogeneously grouped, only in reading and composite score, for both rapid and non-rapid learners, there were no significant differences in personal or social adjustment.

In a study described earlier in this document, Drews (1962) used two self-concept measures. One instrument was the Ability Self-Concept Rating, consisting of a single question asking the student to compare his ability with the abilities of his classmates and to rate himself as above average, about average, or below average; the other was the Concept of Self-As-A-Learner Scale, a 20-item instrument developed by Drews from Bell's Index of Adjustment and Values. The Ability Self-Concept Rating was administered both as a pretest and as a posttest; the Concept of Self-As-A-Learner Scale was administered at the end of the study only. On the Ability Self-Concept Rating administered as a pretest, the one significant difference favored slow students in the homogeneous group; on the same instrument administered as a posttest, superior students in the heterogeneous groups and slow students in the homogeneous groups made significantly higher scores on the instrument. On the Concept of Self-As-A-Learner Scale, Drews found that although heterogeneously grouped superior students obtained higher mean scores, the differences were not significant.

Morgenstern (1963), it may be recalled, compared sixth graders who had been in homogeneous classes for a three-year period with a group that had been in heterogeneous groups over the same length of time. In addition to an achievement test and the California Test of Mental Maturity, she administered the California Test of Personality and Thinking About Yourself. As with achievement, ability grouping did not seem to result in a significantly better personal-social adjustment than did heterogeneous grouping. For students of average IQ, the better personal-social adjustment was found for those grouped heterogeneously.

In a study of homogeneously and heterogeneously grouped students of below-average ability in grade 7 and 8 of two Minnesota junior high schools, Torgelson (1963) administered the Mooney Problem Check List in addition to measures of achievement. On the Check List there was only one significant difference--from beginning to end of year the homogeneous group had a greater decrease than did the heterogeneous group in problems concerned with Home and Family. There were no significant differences between the two groups on sociogram results or in satisfaction with the classroom situation. Torgelson concluded that homogeneous grouping for below-average high school students was not superior to heterogeneous grouping.

Wilcox (1963) studied 1,157 eighth-grade students in 16 schools in five central New York State counties to determine the multiple effects of grouping upon the growth and behavior of junior high school students. The schools were selected to reflect wide variations in grouping practice; the independent variable used was degree of homogeneity of grouping by mental age in the several schools. In addition to instruments designed to measure mental ability, level of achievement, and critical thinking ability, Wilcox used the Maslow Security-Insecurity Inventory, a specially developed Inventory of Attitudes toward Junior High School, and an adaptation of the Ohio Social Distance Scale. He found that, for the total group, self-concept was unrelated to grouping; but for groups in the category below 90 IQ,

there was a more positive self-concept with homogeneous grouping. There were no significant differences in attitude toward school when the total population was examined; but for students with IQ's in the lower categories below 105, it was more positive under homogeneous grouping, and for students of high socioeconomic status who had IQ's 105 or higher, it was poorer under homogeneous than under heterogeneous grouping. Wilcox concluded that, in the absence of curricular differentiation, homogeneous grouping has a significant positive effect upon the attitudes of low normal and low ability students toward self, school, and peers and a significant negative effect upon the attitudes toward self, school, and peers of high ability students from upper socioeconomic homes.

Adkison (1964) studied attitudes about self and group through the use of a questionnaire she developed, and administered in October and again in May to students in grades 3 through 6 in four schools, two at upper-lower and two at upper-middle socioeconomic levels. At each socioeconomic level, the usual heterogeneous grouping was used in one school; homogeneous high and low ability groups, based upon test scores and teachers' judgment, were used in the other. Her findings indicated that low ability students manifested less positive attitudes than high ability groups. The difference was greater with homogeneous groups than with heterogeneous classes, and greater at the upper-middle socioeconomic level than at the upper-lower socioeconomic level. Teachers in homogeneously grouped schools tended to favor such grouping, 44 percent to 31 percent; all who opposed homogeneous grouping were teachers of low ability classes. Adkison concluded that "Homogeneous grouping . . . appeared to be detrimental to those in low status groups and to have a positive effect on those in high status groups. . . . The evidence supports the concept that decisions to separate children through formal grouping patterns should include the question of values."

Bacher (1964) studied 60 slow learners in grades 6 through 8 in a New Jersey suburban school system. Thirty of the pupils were in two special classes, which served as the experimental group; 30 were in regular classes, which served as the control group. The Columbia Classroom Social Distance Scale and the Davidson-Lang Check List of 35 Trait Names were given at the end of the year, and a standardized reading test was given at both the beginning and the end of the year. Bacher found no experimental-control differences in self-concept or reading growth. However, social adjustment of the special class slow learners was significantly more positive than that of the slow learners in regular classes. From this study, Bacher inferred that there is greater acceptance of peers by peers among slow learners in a special class than among slow learners in a regular class.

Deitrich (1964) made a comparison of the sociometric patterns of sixth-grade students in two school systems, one of which used ability grouping and the other, heterogeneous grouping. He found that no appreciable differences existed in the selection of friends between ability-grouped classes and heterogeneously grouped classes, i.e., that ability grouping did not necessarily limit a child in his friend relationships. A strong tendency toward the "bright" selecting the "bright" and the "dull" selecting the "dull" as friends was noted; this was especially true when mutual friendships were involved. He also found that students do not necessarily choose bright students for help with difficult lessons, nor do they always choose a close friend for such help. Deitrich's study indicates

that there are no appreciable differences discernible in the sociometric patterns of sixth-grade pupils who are grouped either heterogeneously or homogeneously.

Dyson (1965) studied two seventh-grade populations similar with respect to age, intelligence, academic achievement, school grades earned, the school environment which they experienced, and the socioeconomic levels of the communities in which they lived. The populations differed in the manner in which they were grouped for instruction. One group was instructed in a school in which students were assigned to classes heterogeneously; the other group, in a school which made a definite attempt to place learners in class sections that were homogeneous with regard to academic learning ability, IQ scores, achievement test scores, evaluations by sixth-grade teachers in the areas of reading and arithmetic, and the principal's evaluation of standing in class. The heterogeneously grouped students numbered 323; the homogeneously grouped, 244. Each of the groups responded to two instruments: the Index of Adjustment and Values, which yields an index of acceptance of self, and the Word Rating List, designed to yield an index of the more specific academic self-concepts. Dyson found that neither the patterns obtained when acceptance-of-self reports were compared with how students were grouped nor those obtained when academic self-concept reports were compared with how students were grouped varied from those to be expected as a result of random variation. He also found that while high achievers did not report significantly different patterns of acceptance of self from those of low achievers either in homogeneous or heterogeneous groupings, they reported significantly different patterns of academic self-concept from low achievers in both heterogeneous and homogeneous grouping situations. Dyson concluded that ability grouping alone did not appear to have a significant effect on either reports of acceptance of self or academic self-concept.

Zweibelson et al. (1965) studied the attitudes and motivation of approximately 360 eighth- and ninth-grade students assigned to three ability "tracks." An attitude survey with seven scores and a motivation inventory were administered before and after exposure to a program of team teaching. Contrary to expectations, the pretesting showed the brighter students in the high ability groups tending to have significantly lower motivation scores than students in the lower ability groups. Students in the high ability groups also tended to have more negative attitudes toward group and school. There was little change in these basic relationships after exposure to the team teaching program; there was, however, at this point a significant positive relationship between the total attitude score and the motivation score not present originally. Zweibelson suggested that ability grouping may create more tension or pressure for the more able student, and that negative attitudes and lower motivation are possible consequences of this.

In the longitudinal study described earlier, Borg (1966) examined a number of non-cognitive variables at various grade levels in addition to achievement: sociometric choices, student attitudes, student problems, self-concept, and personality. During the four years covered by the study, he administered many different non-cognitive measures to different groups at different times. In reporting his study, Borg indicated that the net effect of ability grouping on affective development was probably harmful to at least some of the students

educated under such a system; and that where ability grouping showed any advantage over random grouping, the advantage was usually a slight one. In ability-grouped classrooms at the elementary school level, superior students showed a significant loss in sociometric status while average and slow students made gains in status. At the junior high level, ability grouping was consistently related to fewer problems. Attitude toward peers was found to be consistently related to ability in the random-grouped classrooms while no such relationship was found in the ability-grouped classes. At all levels and for all samples, ability grouping was generally associated with less favorable self-concept scores. With respect to level of aspiration, Borg found no significant differences for students at the same ability levels in his randomly grouped and heterogeneously grouped samples; neither did he find that ability grouping led to a greater feeling of belonging on the part of students at any ability level, but that, instead, it provided a less favorable climate. His personality measures showed that the two grouping treatments did not affect differentially such personality variables as poise, ascendancy, and self-assurance, except in the case of students of average ability, where the random group showed a tendency toward more favorable scores. The Borg data suggest that the method of grouping students is not a uniformly significant factor in the feelings either of superiority or inferiority among elementary and junior high school students. The fact that self-concepts are lower for all groups at all levels and that Borg himself questions whether any small advantages to some compensate for the harmful effects on others, leads us to interpret his findings in this area as essentially negative.

Borg and Pepich (1966) conducted a controlled study of slow-learning tenth graders (IQ between 70 and 90) in a Salt Lake City high school. Students were matched for social class factors and grouped in English classes. Two different classes were studied in two different years; tests were administered at the beginning and end of each school year. The homogeneous grouping resulted in more class participation and more quality contributions. No significant differences were found between groups in either self-concept or attitudes; the only difference between groups was that the number of unexplained absences was significantly higher in homogeneously grouped classes. The authors concluded that the advantages of the more comfortable competition provided in homogeneous groups were outweighed by the disadvantages of the low-group label.

As part of their comprehensive study of the effects of ability grouping, Goldberg, Passow, and Justman (1966) reported student appraisals of their present status and their ideal or wished-for status on a variety of personal characteristics and abilities, as well as on academic expectations and satisfactions. The two instruments used were I Guess My Score and an adapted Index of Adjustment and Values. Although the presence of both gifted and slow students had statistically significant effects on the self-attitudes of the other ability levels, the results were inconsistent. The presence of gifted children tended to result in improved self-attitudes for brighter students and in less positive self-appraisals for slower students, but had little effect on average students. The effects of the presence of slow students varied from one area of assessment to another and also from one ability level to another; the presence of such students was associated with higher expectations of academic success held by the very bright and average students, but there was lower success expectation on

the part of gifted students. Little support was found for the notion that narrow-range classes are associated with negative effects on self-concept, aspirations, attitude toward school, and other non-intellectual factors. In general, the effects of narrowing the range or separating the extreme levels was to raise the self-assessments of the slow students, lower the initially high self-ratings of the gifted, and leave students at the intermediate levels largely unaffected. The slow students also showed greater gains in their "ideal image" when the gifted were absent than when they were present. While grouping appeared to have no negative effects on the self-concepts and school attitudes of students in this study, it must be noted that largely because of the requirement that each participating school have at least four entering-fifth graders with IQ's of 130 or higher, the schools included in the sample were almost all located in predominantly middle-class sections of New York City and that their populations were, as a result, relatively homogeneous with regard to social class; furthermore, the low ability group was of low-average rather than low intelligence and included few students with IQ's below 90. Even for this select population the authors conclude cautiously, thus: "Ability grouping is inherently neither good nor bad, it is neutral. Its value depends upon the way in which it is used. Where it is used without close examination of the specific learning needs of various pupils, and without the recognition that it must follow the demands of carefully planned variations in curriculum, grouping can be, at best, ineffective; at worst, harmful."

Olavarri (1967) studied the relative merits of heterogeneous and homogeneous grouping in terms of the students' self-concepts under these two arrangements in families. The Concept of Self-As-A-Learner Scale was used to secure the responses of ninth- and eleventh-grade students concerning how they felt after two years of homogeneous or heterogeneous grouping. Olavarri found that lower ability groups consistently indicated better feelings of self-worth in the homogeneous setting than in the heterogeneous one, while the top ability group responses showed only a slight favoring of the grouped setting. Olavarri concluded that "Apparently the stigma of group labeling was readily offset by the classroom atmosphere and process." The percentage of "successful grades" was significantly higher in lower ability English classes than in the heterogeneous classes, while the reverse was true for the top groups.

Willcutt (1967) attempted to find a practical way of handling individual differences in the junior high school mathematics program. The entire seventh grade, 240 students, of a midwestern junior high school was involved. Fifty percent of the students were assigned to experimental classes--one review level (low), two standard (average), and one in depth (high)--and fifty percent to the control group. The instructional program was one whereby students were continuously regrouped during the year on the basis of their proficiency in each of the eight different mathematics topics studied. Of the 120 students in the experimental group, only seven remained in the "in depth" class throughout the year and only six in the review class. Pretests and posttests in arithmetic were administered, along with a questionnaire designed to test changes in attitudes. While there were no significant differences in arithmetic achievement between ability-grouped and heterogeneously grouped classes, the flexible ability grouping did result in significant attitudinal changes favoring the experimental group.

Sarthory (1968) studied sixth grade students from six schools in a large metropolitan area in the Southwest. Three schools used heterogeneous grouping, and three used two homogeneous groups, one above and one below the school's median IQ. Varying proportions of Anglo- and Spanish-American students attended the schools. Self-concept was measured by the Sense of Personal Worth Scale of the California Test of Personality; intercultural attitudes were measured by a semantic differential test; occupational aspirations were measured by the Haller Occupational Aspiration Scale; and educational aspirations were assessed by the use of a five-point scale devised by Sarthory. The major findings were that "An ability group cannot be considered as a reference group. Rather, self-concept, intercultural attitudes, and aspirations appear to be based on one's membership in other social groups, particularly the family and socioeconomic status." According to Sarthory, grouping did not significantly affect these variables except for occupational aspirations: the grouped students of high IQ had higher aspirations than the ungrouped high IQ students. There were indications in this study that grouping tended to inflate or deflate slightly attitude sets which were grounded mainly in socioeconomic status and IQ considerations, and that intercultural attitudes were based more on socioeconomic status factors than on ethnic factors. Sarthory recommended that ability grouping not be used. He suggested, instead, the use of techniques of individual instruction, formal preschool programs to remove deficiencies, and the establishment of attendance districts to insure no "perpetuation of tensions of the larger society."

A study by Borg and Maxfield (1967) was concerned with the long-range sociometric development of a sample of students first studied at grade 4 (Borg, 1966) and followed through grade 11 in this later project. Sociometric choice measures were obtained on an initial sample of 1,031 fourth-grade students and subsequently on students available from this initial sample at grades 5, 6, 7 and 11. Subsamples of about fifty students who had made the greatest gains and losses in sociometric status since grade 7 were interviewed and administered an autobiographical questionnaire, a self-concept measure, a school attitude measure, and two personality inventories in grade 11. Analysis of the data obtained indicated that the mean sociometric choice scores obtained at grades 7 and 11 by students in ability-grouped and randomly grouped classrooms were not significantly different at any of the ability levels. Differences in sociometric-choice patterns found at lower grade levels in the earlier study were not present at the secondary level. For four groups of students selected on the basis of scores obtained at grades 7 and 11 and identified as the Low-Low group, the High-High group, the Up group, and the Down group, none of the measures obtained in the earlier grades yielded differences sufficiently large or sufficiently consistent to be of any value in predicting future trends in sociometric status of elementary school students.

Good and Brophy (1969) report observational data on treatment of boys and girls in first-grade reading instruction. They found that differential treatment by sex did not occur in the reading period, but at other times when boys' disruptive behavior drew more rebukes. The children did not separate out these areas, considering boys to get generally more negative treatment. In a reworking of the same data, Brophy and Good (1969) found teachers gave more positive reinforcement to those children they judged most able and more negative or unresponsive reactions to those judged less able.

Ability Grouping--British Style

In order to serve the highly selective university system in England (only seven to eight percent of the young people of college age are at the universities), a sorting out process has, over the past half century, resulted in rigidly "streamed," or ability-grouped, primary schools (ages 7+ to 11+), based on reports of infant schools (below age 7+) and internal and external examination, rigidly "streamed" junior schools (ages 11+ to 16), and separate grammar and secondary modern schools (terminal). Only since World War II have comprehensive schools at the secondary school level emerged. In the early 1950's articles criticizing streaming began to appear, and research on the subject began to be published in the late 1950's. In 1967 appeared the Plowden Committee Primary School Report, which recommended unstreaming in infant schools with the hope that it would spread to primary and junior schools. This hope has not as yet been substantially fulfilled; the latest figures show that 58 to 70 percent of the junior schools still practice some form of streaming.

Ogletree (1969) discussed the pro's and con's of streaming and reported on some of the more significant research. The arguments advanced by British school administrators and teachers are strikingly similar to those advanced for and against ability grouping in the United States. Ogletree reported that most of the research conducted in England indicated that students in lower streams possessed a sense of failure resulting in a consistent decline in morale, effort, and attainment. He offered the opinion that even if streaming gave sound and true homogeneous groups, it "ignores the more subtle aspects of the personality and the social aspects of man."

As indicated earlier in this document, few of the research studies concerned with the advantages and disadvantages of streaming have been concerned with academic achievement. Most have been concerned with the effects of streaming on the social adjustment and attitudes of students. Most of these studies suffer from the use of small samples and are, therefore, inconclusive; the best known studies that examine the effects of streaming on non-cognitive aspects show different results. With the research in Great Britain, as with the research in the United States, everyone can find evidence in previous research to support whichever side he takes on this issue.

Rudd (1958) tested the hypothesis that the attainments, attitudes, behavior, and personalities of students taught in a school organization based upon streaming would be influenced by that organization. His experiment involved two groups of 90 students entering the same school at the age of 11 years. The control group was organized into three heterogeneous classes whose membership did not change during the two-year period following entry to the school; the experimental group was organized into three streams and students were transferred between streams after each half-yearly examination. Neither tests of ability nor tests of attitude toward examinations, school lessons, and school life in general yielded significant differences between groups. Samples of classroom behavior revealed that in the group organized

into streams, fewer social contributions were made by students and there was more aggressive behavior and less attention to work. Estimates of personality made by teachers revealed no significant differences between groups while students' self-estimates revealed an extensive, but probably temporary, deterioration in personality following regrouping. No general long-term effects attributable to streaming were discovered.

Cox (1962) investigated the effects that educational streaming practices have on scores on the General and the Test Anxiety Scale for Children. He used an Australian adaptation of both scales, which he administered to a sample of 266 fourth- and fifth-grade children in two schools in Canberra. In each school, the children had been divided into "superior" and "inferior" subgrades on the basis of their academic records in the first three grades of school. Cox found that general anxiety scores were independent of educational practices but test anxiety scores were significantly, and negatively, related to level of subgrade. He also found that test anxiety scores increased with grade.

Willig (1963) investigated the social implications of streaming by academic attainment in the junior school with particular reference to its possible effects on (a) social interaction between children of differing intelligence and socioeconomic status; and (b) differences in social adjustment and social attitudes between children in streamed and unstreamed classes, and such differences between children in "A" (faster) and "B" (slower) streams. Two hundred boys and girls, aged between 9 and 10 years, were drawn from two contrasting social areas. In each area, an "A" class, a "B" class, and an unstreamed class were studied. A sociometric test was administered to determine social interaction between the various criterion groups. The N.F.E.R. Primary Verbal Test 1 was used as a measure of intelligence, and an index of socioeconomic status was provided by grading occupations of parents. Teacher ratings were obtained to determine incidence of maladjustment, and an attempt was made to measure children's social attitudes by means of a sentence completion test. Other measures included a brief questionnaire designed to explore children's attitudes toward streaming. Taken as a whole, the evidence from the sample pointed to the social advantages of heterogeneous grouping as opposed to streaming by academic attainment. Heterogeneous grouping provided greater opportunities for the formation of mutual relationships between children of different intelligence and socioeconomic status levels. In streamed schools cleavage existing between "A" and "B" streams operated to force the more intelligent "B" class children of intermediate socioeconomic status to associate only with their intellectual and social peers, or with children in lower intelligence and social class groups. There was a tendency for children in unstreamed classes to be superior in social adjustment, as defined by Stott's Six Adverse Adjustment Pointers scale, a relatively crude instrument but one which successfully differentiated between the criterion groups at the 5 percent level of significance. It was also found that in streamed schools "A" class children tended to be superior in measured social adjustment and

socioeconomic status to those in the "B" class. Since social interaction between streams was very limited, "B" class students were prevented from associating with the "better adjusted" "A" class children, who were more likely to conform to a generally accepted system of values. Finally, it was shown that children in streamed schools were fully aware of the advantages associated with "A" class status and of the inferior position of the "B" class in the school hierarchy.

Kellmer-Pringle and Cox (1963) studied 235 children who comprised the entire fourth year in two junior schools in the Midlands. One school was organized in a mainly adult-directed traditional form in which competition, streaming, and class teaching were emphasized. The other school maintained a child-centered progressive regime in which cooperation and the realization of each individual's potentiality was emphasized; in this school, neither streaming nor group tests of any kind were used until the last year in the school. The headmasters of both schools were convinced of the soundness of their approaches and both gave positive and strong support to the staff; each was reportedly dedicated to the welfare of the pupils. On both the General and the Test Anxiety Scale for Children, children in the unstreamed, child-centered, progressive school received significantly higher mean scores (less anxiety) than those in the streamed, adult-directed, traditional school.

Levy, Gooch, and Kellmer-Pringle (1968) carried on a longitudinal study of the relationship between anxiety and streaming in two junior schools, one (School T) a traditional school with streaming throughout and one (School P) a "progressive" school with no streaming until the fourth grade. One hundred eighty-one boys and girls were involved. The General and the Test Anxiety Scale for Children were administered on three equally spaced occasions over a 12-month period. The 11+ examinations* were taken between the second and third testing occasions. Although in some cases GA (general anxiety) and TA (test anxiety) scores yielded parallel findings, differences in school regime and interactions with this factor affected GA scores generally, whereas TA scores showed different relationships with streams on different testing occasions. In School P, GA was found in the lower streams, while in School T the lower stream had the highest mean (less anxiety); these results were broadly true for each testing occasion. The lower streams tended to show more TA but this tendency differed in strength from one testing to the next. In School P, both scores fell on the second testing, but on the third occasion GA remained high whereas TA showed a fall. The investigators suspected that the onset of streaming and the coming of the 11+ examination aroused previously

*The 11+ examination was for a number of years administered universally in Great Britain at the end of the junior school to determine eligibility for secondary school education in the grammar school (academic) or the secondary modern school (terminal). While it is still widely used, it is not as popular as it once was. Critics maintain that it sorts too early and too permanently for many children.

unexperienced anxieties in School P. The passing of the 11+ examination by the third testing might then be supposed to allow TA to fall, even in School P, while GA remained high in that school as a function of the continuing and widespread social effects of streaming.

Griffin (1969) studied 586 children at age 14+ in three grammar, three comprehensive and six secondary modern schools. No systematic differences in educational attainment were found. Children in the comprehensive schools recorded better attitudes toward school; boys and girls in comprehensive schools, at each level of ability, expressed the wish to stay at school longer than did their counterparts in grammar and secondary modern schools although the differences were not significant at the 5 percent level. For children of average and below average ability,the comprehensive schools appeared to provide a more stimulating environment than did the secondary modern schools. If the grammar schools are considered to be upper level and secondary modern schools to be lower level, both homogeneously organized, and comprehensive schools to be heterogeneously organized, this study presents results that are similar to those being reported for a great many studies in the United States for homogeneous versus heterogeneous grouping.

Under the sponsorship of the National Foundation for Educational Research in England and Wales (N.F.E.R.), Bouri and Barker Lunn (1969) made a study of the effects of different types of school organization on student achievement and behavior in 28 junior schools having four classes or fewer. The two main forms of organization were the Traditional Standard method, approaching the homogeneous, involving rough allocation of children to classes according to age but with double promotion of the more able students and retention of the less able, and the According-to-Age, or more heterogeneous method, which adheres strictly to the criterion of age (in months) in the assignment of students. In schools with fewer than four classes, it is necessary to split a year-group of students and put more than one year-group in a class even in According-to-Age schools. Ninety-four teachers and 2,822 students were involved in the study. The two halves of the sample matched satisfactorily on nine out of ten criteria; suitable adjustments were made for the tenth criterion, father's occupation. Teacher ratings and sociometric data revealed no differences in total maladjustment ratings, although on individual traits certain differences emerged. For example, students from all social classes in Traditional Standard schools were considered by their teachers to be more prone to bullying and fighting, and students of the upper socioeconomic group in these schools were rated as more disobedient than their According-to-Age counterparts. On the other hand, students in lower socioeconomic groups in According-to-Age schools were considered more withdrawn and less pleasant to have in class. On the basis of sociometric data, classes in According-to-Age schools had a warmer and more friendly atmosphere.

The larger study conducted by Barker Lunn (1970) under the sponsorship of N.F.E.R. is easily the most extensive ever conducted to examine the effects of streaming and non-streaming on the personality and social and intellectual development of junior school students. A major part of the research was concerned with the follow-up, through their junior school

course, of approximately 5,500 children in 72 junior schools, 36 streamed and 36 unstreamed. The students were tested initially at age 7, in 1964, and then annually until 1967, when they were in their final junior school year. The measurement instruments were tests and questionnaires designed to assess performance and attitudes in nine different areas: (1) attainment in reading, English, and mathematics; (2) verbal and non-verbal reasoning; (3) creativity, or divergent thinking; (4) interests; (5) school-related attitudes; (6) personality; (7) sociometric status; (8) participation in school activities; and (9) occupational aspirations. Information was also obtained on teachers' attitudes toward streaming and other educational matters on their classroom practices and teaching methods. In addition, a limited study was made of parents' attitudes.

One of the most important findings concerned the role of the teacher. Teachers within streamed schools were more united with respect to both their views on educational matters and their teaching methods; in non-streamed schools there was a wide divergence of opinion. About half the teachers in non-streamed schools held attitudes more typical of teachers in streamed schools; this group of teachers created a "streamed" atmosphere within their non-streamed classes, their teaching methods and attitudes tending to reflect the "knowledge-centered" pattern found in streamed schools rather than the "child-centered" pattern found in the non-streamed school. Because this could easily result in modifying the true effects of an educational policy of non-streaming, all analyses were carried out in terms of two teacher-types: Type 1 held attitudes and used teaching methods typical of non-streamed schools and Type 2 was typical of streamed schools.

The children's academic performance, in the main, was unaffected by their school's organization or their teacher's attitude toward streaming, although the attainment of children who were promoted or demoted were clearly affected, the one favorably and the other unfavorably. In general, neither school organization nor teacher-type had much effect on the social, emotional, or attitudinal development of children of above average ability, but they did affect strongly those of average and below average ability. Children of average ability were particularly influenced by teacher-type in the development of their teacher-student relationship and academic self-image. In these two areas, students who were taught by "typical non-streamers" in non-streamed schools were better off than their counterparts in streamed schools; students taught by "typical streamers" in non-streamed schools held the poorest attitudes. Boys of below average ability also had the most favorable teacher-student relationship with typical non-streamed teachers in non-streamed schools; but more boys of below average ability had a good academic self-image in streamed schools. In the development of certain school-related attitudes--attitude to class, "other image" of class, and motivation to do well in school--children of average and below average ability did better in non-streamed schools than children in streamed schools.

The number of streams in streamed schools appeared to be important. Although students in A-streams tended to improve and those in lower streams to deteriorate in their attitudes, the effect was more pronounced in the bottom streams of three- or four-stream schools.

Children in both streamed and non-streamed schools taught by teachers of either type tended to choose other children of similar ability and social class as friends, although there were a greater number of mixed friendships in non-streamed classes. There was little difference in social popularity of children between those in streamed schools and those taught by "typical non-streamers" in non-streamed schools; however, more children of below average ability taught by "typical streamers" in non-streamed schools were friendless or neglected by other children. More children in non-streamed schools participated in school activities; but in both kinds of schools, especially the streamed schools, bright children and children from the higher social classes tended to be more active.

Although parents' educational aspirations for their children appeared to be influenced by the type of school attended, and in streamed school by the stream-level, this was not true of the children's own occupational aspirations. Whether the desired occupation was based upon fantasy or otherwise, there was little difference between the choices of children in streamed and unstreamed schools. The aspirations of the boys seemed to be much more unrealistic than those of girls and ability had less effect on their choice.

Before attempting to summarize the evidence on the impact of ability grouping on the affective development of children on the present scene, a number of observations should be noted. First, studies of the impact of ability grouping on affective development are a more recent phenomenon than studies of impact on scholastic achievement. The studies in the 1920's and 1930's were concerned almost exclusively with the impact on achievement; the earliest study reviewed in the present section on impact on affective development is dated 1948. Second, many of the earliest studies--notably those by Drews, Goldberg *et al.*--were concerned primarily with delineating the impact of ability grouping on "gifted" students in the period after Sputnik when public concern was concentrated on cultivating high competence in mathematics and science, specifically stressed in the National *Defense* Education Act of 1958. The wording of conclusions of these studies points to concern with the affective development of the gifted when singled out for academic excellence and special opportunity; lower achieving groups are treated primarily as the norm group, the great remainder; comparisons are often with only the relatively low, around IQ 100. Third, as with studies of impact on achievement, the earlier studies show more benefits to the low achievers than now when the low achievers and the high achievers have ethnic and socioeconomic overtones.

On the current scene, then, the impact of ability grouping on the affective development of children is to build (inflate?) the egos of the high groups and reduce the self-esteem of average and low groups in the total school population. A new dimension of interpretation has been emphasized chiefly in the British studies of "streaming" where teacher attitude toward achievement is shown to have marked effect. In particular, teachers who bear attitudes of almost exclusive emphasis on academic achievement to the neglect of personal development exercise an especially pernicious influence on low achieving children in heterogeneous classes where the differences are widest.

ABILITY GROUPING AND SEPARATION: ETHNIC AND SOCIOECONOMIC

In the two previous sections, it has been shown that ability grouping has unfavorable effects on the scholastic achievement and the affective development of students placed in low groups, without redeeming benefits to match. To the extent that minority children are overrepresented in low ability groups, then, they are being made to suffer the unfavorable effects of ability grouping. The mechanism by which this placement in low groups operates to the disadvantage of minority children has been illuminated by the McPartland study, previously cited*, in which it was shown that black children did better, other things being equal, in classes in which the majority of their classmates were of the dominant white group, stimulating learning of the minority children by their more effective response to the teaching-learning situation in the classroom.

In this section, the evidence will be marshalled which shows how sharply the minority children are separated from this stimulation by assignment to low, predominantly non-white classes in schools whose total student populations have been desegregated.

The Special Problem of Metropolitan Areas

First, it should be noted that the issue of desegregation and then resegregation by ability grouping is dead and meaningless in situations where inmigration of blacks and outmigration of whites to suburbs or private schools has already reached a point where the total local school population is predominantly black. The difficulties faced by a large metropolitan system's efforts to desegregate were examined in a study by Walker, Stinchcombe, and McDill (1967), who studied school desegregation in Baltimore. These writers found that although both the Baltimore City system and the Baltimore County system have made some progress toward desegregation within each of the systems, when the two systems are considered as a single metropolitan system, no progress at all has been made. They point out that this is because, while segregation within the political boundaries has declined in importance, the county boundary has become the most crucial segregating influence in the metropolitan area; and unless integration can take place across the city-suburban boundary, neither school system, by itself, will be able to effect any appreciable amount of

*See page 22 in this document.

desegreation. They also point out the importance of private and parochial schools in maintaining segregation. Even though concerted efforts might decrease segregation in the public schools, this would have relatively little effect because a very large part of the white school population who might go to school with blacks are not subject to public policy because they attend private and parochial schools.

The progress that has been made so far in the city of Baltimore has been made entirely by introducing blacks into previously segregated white schools; there has been virtually no introducing of whites into formerly all-black schools. Also, the fact that some schools which were previously desegregated have tended to become nearly all black is an indication that the number of predominantly black schools never declines; it always increases. The problem of resegregation has become a factor in the Baltimore schools. The only kind of desegregation that has apparently been implemented in Baltimore has been almost exactly equaled in recent years by a compensating number of schools which became segregated. In the city of Baltimore, there are very few schools left which are still segregated white. These writers point out that, within a few years, it will be impossible for any city policy to achieve desegregation because there will be no more segregated whites to attend schools with blacks in an integrated environment. All of the above forces operate more strongly on the elementary level than on the secondary level; that is, more blacks go to school with whites in secondary schools than in elementary schools. Thus, desegregation progress has been more substantial and longer lasting in secondary schools.

In Baltimore, as elsewhere, the fundamental causative factor for segregation in the schools is the segregated pattern of housing within predominantly black or predominantly white neighborhoods. The elementary schools are almost exactly as segregated as are the neighborhoods in the metropolitan area of Baltimore. Senior high schools are considerably less segregated than the neighborhoods. This is an important aspect of the problem. Whatever influence the public school has on the level of segregation of social life in the city and county of Baltimore, it is more in the direction of desegregation than is true of neighborhoods.

One of the ideas examined in Baltimore is the notion of the "tipping point," that is, the proportion of blacks in a school beyond which whites will leave. The notion of the "tipping point" has been used in the city of Atlanta as an explanation of the tendency for schools which were all white at one time and then were desegregated to later become all black. According to the Baltimore study, the "tipping point" notion does not have validity in Baltimore. Instead of the "tipping point" idea, what is referred to is a demographic pressure in which an increasing black school population pushes about equally on all schools near enough to black neighborhoods for the children to go there. In the Baltimore situation, the fundamental aspect of neighborhood segregation is the differential net migration. As a black moves out of a desegregated neighborhood, he tends to be replaced by another black; however, when a white moves out of a desegregated neighborhood, he tends to be replaced by a black. The net migration, therefore, of whites into the metropolitan area takes place almost entirely in the suburbs, while the net migration of blacks takes place almost entirely by movement into the city. Differential net migration, therefore, constantly increases the blackness of inner city schools.

Viewed as a national problem, the problem posed by the Baltimore situation must be considered typical of virtually every large metropolitan area. The present situation there could be made to confer the benefits of desegregation on minority children only if the city and county schools were consolidated into a unitary school system and all private schools were also required to desegregate. What is said hereafter about ability grouping must be presumed to apply only to the situations outside metropolitan areas where predominant majorities are white, and blacks and other minority groups constitute absolute minorities when whole school districts are considered. In metropolitan areas only drastic procedures of consolidating urban and suburban districts, and transportation of many students, would meet the requirement of equal access to educational stimulation for all groups.

Limited Research on Grouping Practices and Separation

As indicated earlier in this document (p. 2), relatively little attention to the consequences of ability grouping with respect to ethnic and socioeconomic separation is evident in the literature. There are a number of possible hypotheses to explain this omission.

One might argue, as has already been pointed out (p. 40), that the question as to the effects of a particular grouping practice on ethnic and socioeconomic separation is relevant only when the particular environment under study is ethnically and socioeconomically integrated; that is, given a community, school district, or school that is overwhelmingly segregated, it makes little sense to study the practical effect of grouping method X in relation to ethnic and socioeconomic differences in children--not that the question of *de facto* segregation is irrelevant or that it should not be of concern to educators and researchers, but that it is not a researchable question in a self-contained, racially isolated environment.

Further, given the degree of correlation between ethnic origin and socioeconomic class and performance on standardized measures of ability and achievement, to be discussed further later in this document, it seems intuitively obvious, almost without the need for research, that a grouping practice that is based on such measures predetermines the placement of a high proportion of non-white and lower socioeconomic class children to the lowest homogeneous ability groups.

Finally, in the most recent examination of research studies addressed to the desegregated environment, Weinberg (1970) noted that in 1966 a Federal official in charge of desegregation enforcement replied to a Congressional inquiry as to the extent of research on desegregation: "The basic problem is there are few researchers that want to work on it for some reason . . . ." (p. 1).

Notwithstanding the lack of scientific interest, it appears that the problem is probably more than a result of a fundamental dilemma in the American system: the isolation of certain ethnic and socioeconomic groups from the

mainstream of a mixed society. Before, however, discussing ther aspects of the problems and before presenting those few studies which document de facto separation in classrooms as a direct consequence of ability grouping, more extensive discussion of the extent of racial isolation is in order.

Racial Isolation in America

As reported by the U. S. National Advisory Commission on Civil Disorders (1968), hereafter NACCD, there were 21.5 million Negroes in American in 1966. Fifty-five percent of this population lived in the South, 69 per cent lived in metropolitan areas, and nearly half lived in 12 major cities. It is critical to note that, for Negroes, inmigration to the cities has come to mean resegregation. According to Racial and Social Class Isolation in the Schools (1969), hereafter RSCIS, prepared by the Division of Research of the New York State Education Department:

> Overall figures on urban centers do not reflect the segregation of Negroes within the cities. Like other immigrants, Negroes, as newcomers to the city, have lived in the oldest sections. . . . Once in the city, the Negro remains a city dweller. Economic limitations and residential restrictions have barred further movement. But, among the rest of the population, the trend for the past 25 years has been from the city to the suburbs. The combination of inmigration of Negroes and outmigration of white city residents has resulted in disproportionate numbers of Negroes in the cities in comparison with their representation in the total population. This disparity is intensified by the Negro birth rate and will become more pronounced. It is predicted that 13 major central cities of the country will be over 50 percent Negro in 1985. (p. 45)

With respect to the national school enrollment statistics, the inmigration of Negroes and outmigration of whites has had serious implications. For example, the NACCD reports that in the 1965-1966 school year, 17 large city school systems in the nation (including seven of the ten largest) had Negro majorities in the elementary schools. In only two of these cities, Newark, New Jersey, and Washington, D.C., did Negroes exceed 50 percent of the general population.

Even more serious is the finding that within a school system, Negro concentration in individual schools tends to be far greater than their proportion in the total enrollment. As reported in RSCIS:

> In 1965, in 75 major central cities, 75 percent of the Negro elementary pupils attended schools that were 90 percent or more Negro, while 83 percent of the white elementary children were in schools that were 91 percent

> or more white. These school systems were in both the North and the South, and the isolation of the Negroes held regardless of the proportion of Negroes in the total system. (Chapter II, p. 46)

These data tend to highlight a principal finding of the U. S. Commission on Civil Rights, reported in Racial Isolation in the Public Schools (1967):

> The causes of racial isolation in the schools are complex. It has its roots in racial discrimination that has been sanctioned and even encouraged by government at all levels. It is perpetuated by the effects of past segregation and racial isolation. It is reinforced by demographic, fiscal, and educational changes taking place in the Nation's metropolitan areas. And it has been compounded by the policies and practices of urban school systems. (p. 17)

As noted in the 1967 report of the U. S. Commission on Civil Rights, the policies and practices within the school system are seldom neutral in effect. Rather, they reduce, positively reinforce, or maintain ethnic and socioeconomic separation in the schools. Recent empirical studies clearly demonstrate how the educational policy of ability grouping tends to reinforce and, therefore, perpetuate ethnic and socioeconomic separation. In each of these studies, research is focused on a critical dimension of instruction: the classroom composition of children. Several of these studies are presented in detail later in this document.

Ethnic and Socioeconomic Status in Relation to Test Performance and School Achievement

Acknowledging that ability grouping as an educational policy is currently widespread and that student performance on standardized tests is frequently used as the criterion for classifying children into ability groups, then evidence bearing on the degree of relationship between ethnic and socioeconomic status and achievement on standardized measures should be examined to determine the extent to which the practice of ability grouping is likely to separate children along ethnic and socioeconomic lines. The following summary does not claim to be an exhaustive presentation of the research bearing on the issue. Rather, it is intended to present some recent reviews of the literature which suggest that there is a clear relationship between ethnic and socioeconomic status and school achievement as measured by standardized tests and to discuss the conclusions of a few of the most significant research studies.

If there is a paucity of research concerned with the relationship between ability grouping and ethnic and socioeconomic separation, there is no lack of studies concerned with ethnic origin and socioeconomic level in relation to performance on standardized tests. Numerous studies have been conducted on

the relative performance of various ethnic and socioeconomic groups at the elementary, junior high, and senior high school levels. In all, the studies have used a wide variety of tests and measuring devices of school performance ranging from standardized ability and achievement tests, school grades, and teacher ratings, to highest school grade attained and average age for grade level.

Hubert Coleman, writing in 1940, was critical of studies done earlier. In his words:

> A review of earlier studies gives an inadequate and fragmentary picture of the relationship between socioeconomic status and such factors as intelligence, achievement, and personality adjustment. The studies show limitations such as small number of cases, lack of geographic sampling, questionable methods in the measurement of socioeconomic status and intelligence, incidental treatment of the socioeconomic factor, and homogeneous groups with respect to socioeconomic status.

Coleman himself (1940) studied data made available to him by the Advisory Committee of the Coordinated Studies in Education, Incorporated, on 4,784 junior high school students representing high, middle and low socioeconomic levels as determined by a rating scale based on the Sims Socio-Economic Score Card. IQ's were determined by scores on the Kuhlmann-Anderson Intelligence Tests and level of achievement by scores on the Unit Scales of Attainment battery. Coleman found that differences in IQ favored the high socioeconomic group for boys and girls in each grade, with the median IQ falling between the two lower groups and tending to be closer to the lowest group. He also found a definite relationship between socioeconomic status and achievement favoring the high socioeconomic group. Coleman suggested that while his study showed a close relationship among socioeconomic status, achievement, and intelligence, it was not possible to say whether achievement is a result of socioeconomic status or intelligence or to say that intelligence determines socioeconomic status or that socioeconomic status determines intelligence.

Goldberg (1963) reviewed significant changes in recent decades that have created urgent problems for urban school systems. She also discussed the findings concerning achievement and motivation, with particular reference to Negro and Puerto Rican students. Claiming that, as a general rule, Negro children from low-income families achieved less well in schools than did comparable white children, she asked, "What accounts for the consistently lower academic status of children from disadvantaged ethnic groups, especially the Negroes, than of children from lower-class white families living in the Northern cities?"

Dreger and Miller (1964) in a review of studies comparing Negroes and whites published between 1943 and 1958, stated that Negroes by and large scored lower on both traditional and so-called culture-fair tests of intellectual functions, but they noted that Negroes averaged well within the normal IQ range for Whites.

Goldstein (1967), who presented an annotated bibliography of 80 studies, made from 1938 to 1965, concerned with the education of urban youth of low income, wrote:

> It should come as no surprise to the informed reader that, by every conceivable measure, children of low-income families do not do so well in school as children from more affluent ones. The evidence has been presented in full and dramatic detail for the essentially white populations. . . ; for the essentially Negro population. . . ; for the mixed population . . . ; and for cities in general.

Several sources suggest that social class status may have a greater influence on achievement than does intellectual ability as measured by standardized tests. McCandless (1967) summarized the data on the relative contributions of social status and intellectual ability to achievement and concluded:

> From the intelligence test differences between social classes, we would expect differences in school progress, middle- and upper-class children being expected to do better school work than lower-class children. The actual differences in academic achievement between social classes are even more dramatic than the differences in intellectual level. On the whole, lower-class children achieve less well in school than their intelligence tests predict they will, whereas middle- and upper-class children approach their academic potential more closely. (p. 317)

Mostof the research studies of the relationship between ethnic and socio-economic status and test performance have resulted in findings similar to those already cited. Several additional studies of significance are summarized below.

Kennedy, Van De Riet, and White (1963) studied 1800 Negro elementary school children in the Southeastern United States to provide data on intelligence and achievement variables. The Stanford Binet was used to measure IQ, the California Achievement Test to measure achievement, and demographic data not specified to measure socioeconomic level. The study resulted in the following conclusions: With respect to intelligence, the Negro children had a mean IQ of 80.7, but IQ was negatively correlated with age. IQ was highly correlated with socioeconomic levels though the differences were small between urban and rural residents. There was a significant difference in the mean levels of achievement test scores between the sample and the standardization group, and this difference increased with age. Achievement also correlated with socio-economic level.

Deutsch and Brown (1964) explored intelligence test differences between 543 Negro and white first and fifth graders in different social classes, with particular focus on the lower class. The presence or absence of the father in the home was examined, and whether the child had had organized preschool experience. Social class was measured by a scale derived from prestige ratings of occuptions as well as education of main breadwinners. IQ was measured by

the Lorge-Thorndike Intelligence Test. Differences between scores of Negro and white children were significant and were equally strong at all class levels. Negro children at each socioeconomic level scored lower than white children and Negro/white differences increased at each higher socioeconomic level.

With respect to secondary school, Goldstein (1967) noted a body of data, from Project Talent (Flanagan et al., 1964). Examination of these data in terms of socioeconomic differences tends to confirm the thesis that socio-economic status is related to achievement. In this study, a two-day battery of tests and questionnaires was administered to 440,000 students in 1,353 high schools, "carefully selected to be representative of American secondary schools." The data indicated that, on the basis of a measure of general academic aptitude, males below the median were twice as likely as males in the top 20 per cent to come from families possessing "only the necessities of life." Moreover, while over half of those in the lower 50 per cent came from blue-collar families, less than one third of those in the top 10 per cent did so. Rather, about 57 per cent of the latter group came from white-collar families, while only 15 per cent of the students in the lower 10 per cent did.

In addition, Project Talent schools were classified into two relatively homogeneous middle- and low-income groups. One such group consisted of 27 schools that served predominantly middle-income students in New York City, Philadelphia, Detroit, Chicago, and Los Angeles. According to Goldstein, "there was virtually no overlap of the middle two thirds of the two populations, with low-income students consistently below middle-income students in the same school system."

Miner (1968) collected data from the files of 663 high school graduates in a Midwestern city to investigate the relationships between a number of sociological factors, among them social class, family structure, and school achievement, at various periods in the child's academic career. Tests for which scores were available include the California Mental Maturity Test, the Iowa Tests of Basic Skills, and the California Achievement Tests. Secondary school grades were also used. Significant relationships were found between a child's background and his early achievement. For the most part, the differences were small, but they were large enough to account for some of the variance in academic performance. Socio-economic status was found to be positively related to the measures of performance.

In Racial and Social Class Isolation in the Schools (1969), it was concluded that racial differences in achievement are approximately of the same order as the IQ differences between Whites and Negroes. Data from the Coleman Report (1966), based on a test of verbal aptitude, suggest an average difference in IQ of approximately one standard deviation between black and white children at grades 6, 9, and 12 in the Metropolitan Northeast. According to RSCIS, data from these grades also indicate a difference of approximately one standard deviation in the achievement levels of Whites and Negroes of the Metropolitan Northeast. These deviation scores indicate that relative differences in achievement of Negroes and Whites remain constant from grade to grade; grade equivalent scores indicate that these differences grow larger with successive grades. According to RSCIS, the interpretation of Negro-White achievement differences in grade equivalent scores as showing an increasing divergence with years in school is

inappropriate for Negro-White comparisons. The conclusion reached in RSCIS was that the Coleman data, correctly interpreted (in standard deviation units), show that achievement differences between Negroes and Whites do remain relatively constant from year to year.

Unfortunately for the purposes of this research, grade equivalent scores become progressively less meaningful in junior and senior high school; in fact, the decelerating curve of growth on tests of basic skills might spuriously magnify differences expressed in grade scores. However, differences expressed in standard deviation units of white students of a given grade eliminate all opportunity to reflect increases in differences in average performance insofar as variability of individual achievement increases with age and schooling. The fact that grade score equivalents in the middle and upper elementary grades constitute approximately equal units and show progressively increasing differences between blacks and whites, makes safest the interpretation that differences continue to increase, but in a fashion uncertainly represented by grade score equivalents.

Goldstein (1967) observed that although the instances have been few, some studies have come up with contrary findings. For example, Antonovsky and Lerner (1958) found that on the basis of a small class-matched sample of Negro and white students from lower socioeconomic status (complete data were available for 61 Negroes and 54 whites, about equally balanced for sex), the Negroes, despite greater handicaps, did as well academically as the whites, dropped out of school less frequently, and enrolled more often in the College Preparatory program.

Goldberg (1963), in the reference previously cited, cautioned:

> Despite consistent differences in demonstrated intellectual and academic ability. . . there is a great deal of overlapping. In all studies there are some in the one group who resemble the other group far more than their own. And in all comparisons of lower- and middle-class children there is a sizable though smaller proportion of the former who score high on tests, do well in school, plan on advanced education, and have a high degree of similarity to the school performance of middle-class children. Conversely, there are middle-class children whose motivation and performance are poor indeed. (p. 81)

Despite some few exceptions, it appears from the above discussion that, for the majority of the population, ethnic and socioeconomic class variables consistently tend to be associated with school achievement as measured by widely used standardized tests. What does this mean with respect to the placement of children in elementary and secondary schools?

Empirical Consequences of Ability Grouping on Ethnic and Socioeconomic Separation in the Classroom

In view of the high degree of relationship between ethnic and socioeconomic status and performance both on standardized tests and in the classroom, it stands to reason that the use of ability grouping as a strategy for organizing children into classroom units should result in the separation of children along ethnic and socioeconomic lines. While, as has been indicated earlier, few research studies have been directed to separation along those lines, the studies that have been made show that such separation surely does exist, with children from the middle and upper classes found mainly in the middle and upper ability groups and children from the lower classes in the low ability groups.

In Racial and Social Class Isolation in the Schools (1969), several studies are cited which show that grouping on the basis of achievement or aptitude tests leads to ethnic and socioeconomic isolation. Just as there are learning interference factors related to "inferior" schools, the report states, learning interference factors "should also be relevant in schools with grouping policies which result in either social class isolation within schools or combinations of different levels of racial and social class isolation, depending upon the class status of the white student population and proportion of 'integrated' Negroes in the school."

Heathers (1969) in his review of the literature on ability grouping reported only four research studies concerned with the separation that can result from such grouping, none of them done in the United States. Despite the sparseness of research data, however, Heathers wrote:

> It is commonly recognized that low-ability groups in elementary school have a disproportionate number of boys, of children from lower class origins, and of children from minority groups. Ability grouping may thus be, in effect, an agency for maintaining and enhancing caste and class stratification in a society.

In the current search of the literature several studies have been located which support the notion that ability grouping tends to isolate students of one ethnic group or socioeconomic level from another and that this isolation has deleterious effects upon various aspects of the development of students so separated. If, as a growing body of literature indicates, the impact of a school upon individual students is a function of peer interactions--that is to say, that students tend to learn as much from other students as they do from teachers, then these adverse effects can be anticipated.

Mehl (1965) studied 654 students in grades 5 through 8, who had been assigned to classes on the basis of group intelligence test performance from grade 4 on, to determine whether homogeneous grouping is an aspect of school procedure which may reflect, and thus reinform, the social structure

of the community. Social class was determined by Warner's Index of Social Class scale. The same pattern of social class segregation was obvious in all four grades. Although all five social classes were proportionately represented in the two middle-ability groups, in the two top and two bottom groups there were statistically significant differences between the proportion of each social class level in the group and the proportion for the grade as a whole. Segregation was most pronounced in the extreme high and extreme low ability groups. A high relationship was found between measured IQ and achievement; a moderately low relationship was found between IQ and social class and between achievement and social class.

Wilson (1967) in his study of students in Richmond, California, found a marked relationship between the social class composition of schools and student performance. Regardless of their own social class, Richmond students were more likely to perform well in predominantly middle-class than in predominantly lower-class schools. When the relative importance of individual and school social class was assessed for black and white students separately, it was found that the student environment had a stronger relationship to the performance of black students than to that of white students. The performance of white students although strongly related to the social class level of their fellow students, was more closely related to family background than was that of black students.

Wilson also weighed the effects of the social class composition of the school upon the same students over their entire elementary school careers. He found that in the primary grades the influence of individual's social class was of great importance and that the social composition of the school was of little importance. However, over the period of eight years of school, the cumulative effect of the social class composition of the school increased sharply, so that in the eighth grade it was as significant as the individual's social class for student performance.

This pattern was generally the same where student attitudes were concerned, especially with regard to college aspirations and plans. College plans were found to be more frequent for both black and white students in schools with a higher social class level. Black students in schools of lower social class level, even though relatively advantaged, were less likely to attend college than similar students who were in school with a majority of more advantaged students.

In another "study" of the problem, Hobson vs. Hansen (1967), the basic question presented to the Court was whether the District of Columbia Board of Education unconstitutionally depirved the district's Negro and poor public school children of their right to equal educational opportunity with the white and more affluent school children. The case is directly related to the issue under discussion since it was the practical consequence of a track system which gave rise to litigation. Inasmuch as the court decision involves one of the most comprehensive discussions of every major issue introduced in other sections of this document, the relevant evidence presented to the Court will be presented in considerable detail.

The track system used in the Washington, D.C. schools was based completely on ability classification by standardized tests. Accordingly, students at both the elementary and secondary school levels were classified into separate, self-contained curricula or "tracks," ranging from "Basic" for the "slow" student to "Honors" for the gifted. The educational content ranged from the very basic to the very advanced according to track placement. In the elementary and junior high schools, three levels were used: Basic or Special Academic (for "retarded" children), General (for average or above-average students), and Honors (for the gifted). In the senior high schools, a fourth track (Regular) was added for college preparatory training of above-average students.

With regard to the pattern of socioeconomic separation occurring in the schools as a direct result of tracking, evidence submitted to the Court showed that when the high schools were grouped into three levels by median neighborhood income--high ($7,000 to $10,999), middle ($5,000 to $5,999), and low ($3,000 to $4,999)--the correspondence between track placement and income was exact. (Table 3) The economic-level correlations found in high schools were also found, generally, in junior high schools and elementary schools. The Court properly concluded that a student's chance of being selected for one of the higher ability tracks was "directly related" to his socioeconomic background.

Table 3

Per cents of Students in Four Tracks in Washington D.C. High Schools Serving Different Socioeconomic Levels of Neighborhood (1964, 1965)

Median Neighborhood Income	Special	General	Regular	Honors
Over $7,000	0-7.4	7.8-43.7	46.1-80.0	10.2-17.1
$5,000-$7,000	4.7-9.9	39.0-57.7	32.9-49.2	3.2-7.8
Under $5,000	9.8-18.2	54.4-74.5	11.4-33.4	0-3.9

With regard to the pattern of racial separation in the schools, the Court noted that for a majority of District schools and school children race and socioeconomic status were intertwined. The schools serving neighborhoods with income levels of $6,000 or below had Negro enrollments of well over 90 per cent. The only predominantly white senior high school, serving a neighborhood of average income $10,374, had all but eight per cent of the students in Regular and Honors tracks in 1964 and 1965; no other school came close to that. A predominantly Negro school (90 per cent) that was closest served a neighborhood with the third highest income level in the system ($7,650), but had

40 per cent of its students in the lower non-college preparatory tracks. Of the six junior high schools having from 17 to 99 per cent white enrollment in 1964, all six had Honor tracks; at least three of the schools were in the middle-income range. In six other middle-income schools, with student bodies better than 95 per cent Negro, only three had Honors tracks in 1964, and this number dropped to two in 1965.

In reference to the distribution of track offerings in the elementary schools, only 16 per cent of all Negro students were attending schools with Honors programs in 1965. Conversely, 70 per cent of all white students had this advanced curriculum in their schools. This pattern of Honors Track offerings in elementary schools also existed in the junior high schools.

Over and beyond the evidence presented above, the Court made a matter of record further data which illustrated how ability grouping practices result in the ethnic and socioeconomic separation of children. Looking at the racial breakdown of the enrollment in the Special Academic or Basic Track, the Court noted that at both the elementary and junior high school levels the proportions of Negroes enrolled in the lowest track exceeded their proportionate representation in the total student body. On the other hand, the proportion of whites enrolled in the Special Academic track was significantly lower than the proportion of whites in the total school enrollment. It was clear that, as a general rule, in those schools with substantial numbers of both white and Negro students, a significantly higher proportion of Negroes than whites will go into the Special Academic track (for "retarded students").

In summarizing the evidence, it was noted that the track system is by definition a "separative" educational policy, ostensibly according to students' ability level. However, the practical consequence of ability grouping is by its application to separate students largely according to their socioeconomic status and, to a lesser but observable degree, according to their ethnic status.

In recapitulating all the evidence and testimony, the Court pointed out the manner in which the concept and practice of ability grouping structures failure in black and lower socioeconomic class children, perpetuates unlawful de facto discrimination, and generally permeates an entire school system.

The point to be made here, it should be noted, is not to assess intent or blame. The finding is one of fact: that ability grouping produces segregation of students by socioeconomic status and, as a corollary effect, produces segregation by ethnic status. Insofar as such segregation has been shown to reduce stimulation of the low-achieving students to higher educational attainment, the effect of such ability grouping must be deemed to afford less than equal opportunity to the minority ethnic and lower socioeconomic groups.

Very dramatic evidence of how ability grouping based solely on test scores can effect decided ethnic and socioeconomic imbalance in the classroom is given by the following unpublished data made available by a Southern school district which was challenged in Court for its proposal to group black and white children in grades 3 through 8 in multiple sections on the basis of scores on tests in the SPA Achievement Series.

Recommended section assignments for children in grade 5 in five subject matter areas are shown in Table 4.

Table 4

Recommended Section Assignments Based on Battery Test Scores--Grade 5

	Reading		Mathematics		Language Arts		Social Studies		Science	
Section	Black	White	Black	White	Black	White	Black	White	Black	White
A	3	28	3	28	5	26	3	28	3	28
B	4	27	5	26	7	24	4	27	5	26
C	10	21	14	17	12	19	10	21	14	17
D	15	15	15	15	11	19	15	15	15	15
E	23	7	18	12	21	9	23	7	18	12
F	27	3	27	3	26	4	27	3	27	3
TOTAL	82	101	82	101	82	101	82	101	82	101

Reading test scores for grades 3, 4, 6, 7, and 8, shown in Table 5 are typical of the scores in all five subject matter areas for these grades and, consequently, typical of recommended section assignment.

Table 5

Recommended Section Assignments Based on Reading Test Scores
Grades 3, 4, 6, 7, and 8

	Grade 3		Grade 4		Grade 6		Grade 7		Grade 8	
Section	Black	White	Black	White	Black	White	Black	White	Black	White
A	1	29	2	28	4	30	3	31	1	33
B	2	28	3	27	7	26	12	22	14	20
C	13	17	8	22	17	14	12	21	18	15
D	20	10	13	17	14	15	14	18	21	11
E	22	8	21	9	20	8	25	4	23	6
F	22	3	22	5	20	5	22	5	26	1
G	10	1	23	2						
TOTAL	99	96	92	110	82	98	88	101	103	86

After hearing testimony on the total plan for use of this ability grouping for organization of classes in the desegregated schools of the district, the Court ruled against the plan and in favor of a prior heterogeneous grouping plan with special instructional arrangements related to the disabilities being remediated.

Kariger (1962) studied the effect of an ability grouping plan used in the three junior high schools of a Midwestern city of 100,000 on socioeconomic stratification. In this plan, test scores were supplemented by teachers' and prinicipals' judgment in making initial assignments to classroom groups and in making reassignments during the school year to correct for apparent misplacement by original assignment in the light of the subsequent academic performance of the students. Consideration of "teacher grades, study habits, citizenship and industry, social and emotional maturity" were allowed to guide these judgments.

The tracking system called for placing those more than one grade advanced in the high track, those more than one grade retarded in the low track, and those less than one track above or below the norm in the middle group. Reassignments were often required to rectify class size, however.

In keeping with relations found quite uniformly in other studies, assignment to tracks on the basis of standardized test scores above would have resulted in 77 per cent of upper socioeconomic status children in the high track and only 38 per cent of the lower socioeconomic status children in that track. Conversely, only 5 per cent of the upper socioeconomic status children would have fallen in the low track while 26 per cent of the lower socioeconomic status children would have been so classified. However--and this is the thrust of the study--80 per cent of the upper socioeconomic status children whose test scores would have warranted placing them in the high track were actually in that track, while barely 50 per cent (210 of 408) of the lower socioeconomic status children who qualified for high track placement on tests alone were so assigned. Children of the middle socioeconomic group fell into an intermediate position, 65 per cent of those qualified by tests being assigned to the top track.

At the lower end, too few upper socioeconomic status children fell into the bottom track on test scores so comparisons at that level can be made only between middle and lower socioeconomic status children. Again, 37 per cent of middle socioeconomic status children who qualified for the bottom track on test scores alone, were placed in higher sections, while only 15 per cent of lower status children whose test scores would place them in the bottom track were actually placed higher. To summarize, socioeconomic status of children significantly influenced track placement.

Turning now to the practice of reassigning upward children whose classroom performance reflected errors of too low placement initially, Kariger found that only 3.4 per cent of students were affected; but, 70 per cent of all reassignments were to higher classes. However,

93 per cent of changes of upper socioeconomic status children were upward, 68 per cent of middle group children reassigned were raised, and only 61 per cent of the lower group changes were upward. The irony of it all is that the administrators were new to their schools and produced the initial separative socioeconomic effect without any history of prior bias or discrimination against the children based on experience with them.

A study of the Plainfield, New Jersey, school system was conducted by the Institute of Field Studies of Teachers College, Columbia University to determine the practical consequences of the prevailing practices of ability grouping then in use at all grade levels. A 1967 statement of the Plainfield Board of Education expressed its policy in these terms:

> We recognize that within the Plainfield School System there are many different needs and opportunities for class and subject groupings. In order to meet these needs, there may be classes which can now be called racially imbalanced. It is our opinion that it is better to have such classes than not; that these classes should have an objective to prepare for the need for fewer such clases. We also recognize the opportunity for the display of ingenuity and innovation on the part of the staff to minimize any adverse aspects of such racially imbalanced groupings.

The effect of this policy is reflected in Hubbard Junior High School (1968-69) as shown in Tables 6 and 7. In Table 6, the data are for per cents of the two separate ethnic groups in eighth grade to be found in the W (High) track, X (Middle) track, and Y (Low) track in each subject area. Table 7 gives the per cents of total groups in each subject in each track in eighth grade that are black and white, respectively. All figures are to be compared to an overall total of 218 black and 90 white eighth graders, or 70.8 per cent black and 29.2 per cent white. Viewing the data either way, the whites are overrepresented in the top groups and the blacks are predominant in the bottom groups.

Table 6

Percentages of the Hubbard Junior High School, Plainfield, New Jersey Black and White Eighth Grade Students, 1968-69, Enrolled in W, X, and Y Ability Groups by Subject Area

Subject	Race	Group W	Group X	Group Y	Total
English	Black	8.7	48.2	43.1	100.0
	White	58.9	34.4	6.7	100.0
Social Science	Black	10.6	46.8	42.7	100.1
	White	55.6	38.9	5.6	100.1
Mathematics	Black	3.7	56.9	39.4	100.0
	White	42.2	51.1	6.7	100.0
Science	Black	2.8	58.8	38.5	100.1
	White	43.3	50.0	6.7	100.0

Table 7

Percentage Composition of W, X, and Y Ability Groups, Hubbard Junior High School, Plainfield, New Jersey, Eighth Grade, 1968-69 (By Race)

Subject	Group W			Group X			Group Y		
	Black	White	Difference	Black	White	Difference	Black	White	Difference
English	26.4	73.6	47.2	77.2	22.8	54.4	94.0	6.0	88.0
Social Science	31.5	68.5	37.0	74.4	25.6	48.8	94.9	5.1	89.8
Mathematics	17.4	82.6	65.2	72.9	27.1	45.8	93.5	6.5	87.0
Science	13.3	86.7	73.4	73.8	26.2	47.6	93.3	6.7	86.6
TOTAL	22.2	77.9	55.7	74.6	25.4	49.2	93.9	6.1	87.9

The upshot of this survey is significant. After pondering the evidence of ethnic segregation produced, the Board of Education took the following steps toward a more heterogeneous plan:

> To the extent possible, school principals in K-4 buildings have attempted to devise a planned heterogeneous grouping. In the spring, every teacher submits to the building principal a list of pupils in his class, noting whether each child 1) was reading at a high, average, or low level, 2) had been a discipline problem, 3) was Black or white, 4) was a boy or a girl. Using this information, principals attempt to develop self-contained classes composed of a "balanced" representation of children according to sex, race, and achievement, with discipline problems distributed as well.

Thus, the same test data used to produce homogeneous grouping can be used to define and establish heterogeneous groups. It remains to be seen how far and how fast this type of planning is extended to other grade levels.

McPartland (1968), in the complete report of the study mentioned earlier (p. 22), investigated some of the possible ways in which school desegregation might affect secondary school Negro students. His data for 5,075 ninth-grade Negro boys were based on and derived from the mass study by Coleman et al., Equality of Educational Opportunity (1966). McPartland pointed out that in making comparisons between Negro students in situations where the proportion of white students is different, one could look at the proportion of white students enrolled in the school attended by Negro, or at the proportion of white students in the classes attended by a Negro. He used the Coleman data to compare the influence of desegregation at both these levels.

The information collected from students in the Coleman study concerned (a) the students' programs of study, (b) the particular courses in which students were enrolled, and (c) the track levels to which students were assigned in their English classes. It is clear from McPartland's analysis that within schools of similar racial composition the program of study in which a student is enrolled has a strong influence on the chance that he will be in a majority white class. Generally, students enrolled in the College Preparatory program are most likely to be in classes which are more than 50 per cent white. Conversely, students in Vocational, Commercial, or Industrial Arts programs are least likely to have mostly white classmates. McPartland points out that the schools which are exceptions to this generalization are those where only a small fraction of the student body is white. The reason for this is that in contrast to most other schools, "the white students in many of these predominantly black schools are among the poorest students in the school." Therefore, except for predominantly Negro schools with a few white students, the practical consequence of program assignments within schools on the racial composition of a Negro student's classes is the same. Students who tend to achieve in academic areas, as measured by various reading and arithmetic achievement tests, tend to be selected or enrolled in advanced academic programs which tend to have more white classmates than non-academic courses of study.

McPartland presents additional data which highlight the relation between program of study and classroom racial composition. These illustrate that within schools of similar racial composition, black children in mostly white classes are most frequently enrolled in academic courses, and least likely to be taking Vocational, Commercial, Industrial Arts, or Home Economics courses. Says McPartland:

> The most dramatic positive differences with the fewest reversals are for courses which are likely to be part of a college preparatory program rather than some other program: the science and foreign language courses. But even for the course work likely to be required for most students, such as English and mathematics, there is some evidence that enrollment in these subjects is related to the racial composition of a Negro student's classmates. It is with courses such as mathematics and English that separate classes will be organized according to the achievement level of students to be assigned to the class. (p. 99)

Also, with respect to the racial composition of <u>classes</u> as a direct result of tracking or ability grouping, McPartland documents that the largest proportion of the students in the highest track have mostly white classmates. That is, half of all black children in the high English track have more than half white classmates in schools which enroll 50 to 69 per cent whites, while approximately 33 per cent of the Negro students in the middle and lowest tracks are in such classes.

Finally, McPartland goes on to show, as explained on page 22 of this document, that this separation of pupils ethnically has an effect on achievement of the Negro students. Carefully controlling for home background factors, he shows that only when a majority of classmates of black students are from the predominant white group do the Negro students show benefits from desegregation. It is the improved learning of these black students that makes Negro achievement in desegregated schools improve on the average; students in other classes show no improvement and even possibly slight loss.

Mayeske (1970) in the article cited earlier in this document (p. 22) reported further data from the Coleman Report (1966) that are especially pertinent here. In his analysis of the data, Mayeske found a relationship at the first grade level between achievement levels of entering students and the attributes of the schools they attended. Schools with entering students of higher levels of achievement had associated with them teachers who possessed higher verbal skills, who tended to be white, and who expressed a preference for working with high ability students. He found that these relationships with achievement tended to increase at the higher grade levels. The same was true of the relationship of achievement with the students' social background.

Mayeske refers to this phenomenon as the"ecological-functional dilemma." At the beginning of the first grade, students tend to be allocated into schools on the basis of their social backgrounds. Certain relationships, which Mayeske refers to as ecological relationships, are observed between the attributes of the students and their schools. Over time, since students with high social backgrounds benefit more from their schooling, ecology and the school's influence become more and more intertwined so that it becomes increasingly difficult to separate out their independent influences. The schools reflect the deep-seated social problem of ethnic separation which permeates almost every aspect of American life. This basic problem according to Mayeske, in the main is that a person's birth into a particular stratum of society plays a large role in determining where that individual will go and will not go in the scheme of things. The problem is made even more difficult because one's skin color and language habits tend to be associated with one's position within the social structure. If Mayeske's interpretation has any validity, the schools alone cannot rectify the problem, although they can play an ameliorative role; the problem must be attacked on a number of different fronts, such as jobs, housing, schooling, and various other areas characterized by separation and segregation.

Mayeske concludes, as did Coleman, that the schools play an important role in promoting achievement for all students; but, as the schools are currently constituted, students from the higher socioeconomic levels, of whom most are white, benefit more from attending school than students from the lower socioeconomic strata, many of whom are non-white. He suggests that to break these socioeconomic background barriers, innovations that differ radically from past practices might be tried in situations so structured that the results of the innovations can be clearly demonstrated. Some suggested innovations include

more socioeconomically and racially balanced student bodies and teaching staffs, competitive school systems or voucher systems whereby the student and his family can select services from a variety of sources, and concern by real estate people with the improvement of the quality and composition of schools rather than with the maintenance of racially segregated communities in terms of available housing.

Non-Negro Minorities

Many of the educational disabilities which burden Negro Americans are shared by Mexican Americans, Puerto Ricans, and Indian Americans. Weinberg (1970) goes so far as to say that these three minority groups are the most educationally disadvantaged in the United States.

The urban Negro ghetto is reenacted in the Mexican-American neighborhoods in the cities of southern California and the Southwestern states; the Puerto Rican communities in New York and other cities of the Northeast are as isolated from the white communities as is Harlem; and the Indian Americans, especially those living on or near reservations, are the most segregated of all. In recent years, a fourth minority group, the expatriate Cubans in the Southeastern states, especially Florida, have become groups alone.

Belonging to an ethnic minority in the United States and being poor besides creates a common plight for all these people. For Mexican Americans and Puerto Ricans--and, more recently, the Cubans--a "foreign" language has become a barrier to normal educational progress. The exclusive use in most schools of English as the language of instruction, among children understanding this language little or not at all, by teachers not knowing Spanish, has created multiple problems. Add to this the lack of sensitivity on the part of teachers to sociocultural differences in children, and an almost intolerable situation exists in the schools.

Weinberg (1970) devotes an entire chapter in his Desegregation Research: An Appraisal to summarizing research studies of the past 35 years devoted to the exploration of the problems of these minority groups. The research findings are similar to those reported earlier for Negro students and for both black and white students of low socioeconomic status. On the whole, children of the non-Negro minority groups compare unfavorably with middle-class white children with respect to IQ and academic achievement level, segregation has been their usual lot in school, they consider themselves to be inferior to the majority whites, and their educational and occupational aspirations are likely to be low. As with the Negro, in those schools in which ability grouping is practiced, classes almost homogeneous racially have been created. And as with the Negro also, the greater the degree of contact the minority child has with the white man's culture, the higher he scores on educational tests, the greater his progress academically, the more favorable his self concept, and the higher his aspirations.

Carter (1970) has described in detail the history of the educational neglect of Mexican-American children. While there are some exceptions, the majority of Mexican Americans have lower-class status. Even though the children may attend mixed schools, in reality they may be isolated from their Anglo and

middle-class Mexican-American peers. School policy and practice have contributed to this isolation, tending to reinforce the ethnic and social cleavage that exists in the Southwest. The school reflects the community and tends to perpetuate the separation of Mexican and Anglo roles and aspirations.

Special compensatory programs for Mexican-American children are becoming almost universal in Southwestern schools. Compensatory classes requiring attendance for part of the day are most frequently encountered; this kind of program does not isolate the children to an unwarranted degree. When compensatory programs require fulltime attendance, the Mexican-American children are substantially isolated, in essence attending, within an ethnically mixed institution, a subschool from which they cannot break out.

According to Carter, rigid ability grouping, or tracking, in one form or another is widely practiced in Southwestern schools. Appraisal of intellectual capacity and academic achievement, whether by standardized tests or other means, usually determines track assignment. Since Mexican-American children, especially those of low socioeconomic status, tend to fall below school or national norms, they are greatly overrepresented in the lower-ability tracks while the Anglos are overrepresented in the middle- and high-ability tracks. Although a first grader has a better chance to change tracks than a tenth grader, once a student is tracked at any level, movement upward is difficult.

Little research concerning the effects of tracking on the achievement and attitudes of Mexican-American students has been done. Regardless of the effects on achievement, however, Carter contends that the track system adversely affects both teachers' and students' expectations and their subsequent behavior. Since it unduly isolates Mexican-American youth from equal-status interaction with others, it maintains cultural differences and slows down the process of acculturation.

Carter writes that the information collected concerning the practice of tracking in the Washington, D.C., schools at the time of the Hobson vs. Hansen case could equally well describe the practice in most Southwestern schools. To what degree the impact of the Court decision in the Hobson vs. Hansen case may influence Mexican-American organizations to attempt legal recourse to obtain equal educational opportunities for Mexican-American children is a matter of conjecture at the present time.

In response to a request for information about grouping practices based on test scores and school problems they might present for American Indian children, Havighurst (1970) offered this information based on the National Study of American Indian Education:

> . . .most Indian children are in schools where they are in the majority. In these schools, most of which are relatively small, there is seldom any ability grouping.
>
> Another category of Indian student consists of those who live near an Indian reservation but attend. . . a high school that has a majority of non-Indian

students (for example, Cutbank, Montana; Moclips, Washington; Gallup and Albuquerque, New Mexico; Globe, Arizona). In these communities the Indians generally perform below the average of the non-Indian. However, there is not much grouping in these communities, which are generally rather small in their school populations.

A third category consists of Indian students in relatively large urgan centers where the Indians seldom go above 10 per cent in any one school and often are present in less than one per cent proportions. Here there may be some ability grouping based on tests and depending on the policy of the school system. Almost all of the big cities from Chicago on West have these kinds of Indian minorities. Also, you find them in smaller urban centers like Mesa, Arizona; Bell Gardens, Los Angeles; Tucson, Arizona. At the high school level we find that there is some ability grouping based on tests in a number of high schools. Generally, the Indian youngsters tend to be placed in the average or below average ability groups. Still there are usually a few who do well on tests and get placed in the higher ability groups.

It would appear from Havighurst's letter that American Indian children are less generally affected than children of other minority groups by ability grouping practices. Certainly there are no situations in which they are isolated from the white majority as a result of ability grouping. The reader may wish to refer to the study of Maynor discussed briefly on page 21 of this document.

In summary, the reported information about non-Negro minorities is scant, but consonant with the findings for Negro students. The special connotation of "language handicap" for Spanish-speaking or bilingual minorities in the United States could be studied in terms of test results, but is more properly seen in the broader context of pluralistic education, needed respect for minority cultures, and humanitarian concern for all children on an equal basis of acceptance and assistance as well as opportunity.

SUMMARY AND CONCLUDING REMARKS

This second document summarizes, in as readable format as we could devise, the important studies relevant to The Impact of Ability Grouping on School Achievement, Affective Development, Ethnic Separation and Socioeconomic Separation. It is supported in detail by an extensive bibliography of historical and timely references. The reader may expect to find here sufficient discussion of major findings and enough illustrative material to clarify the points made. Careful perusal of the references will allow the reader to fill in the greater detail he may desire at any point without our having slowed other readers not interested in so much detail about that point. On the other hand, we would suggest that what is presented here will be merely supported, clarified or expanded,

but not contradicted in any essential respect by reading the references. Nor do we feel we have omitted relevant references. So far as we could make it, then, this is a summary and guide to the essential truth about this topic.

We are concerned here with schemes of organization of schools into classroom groups on the basis of test results or judgments relative to the ability of students, in such a way as to bring together in instructional groups children of a given age or grade who are most nearly equal in relevant abilities. Grouping and regrouping within the classroom for instruction of those needing assistance in mastering particular bits of skill or information is considered a normal and desirable instructional practice.

Briefly, we find that ability grouping as defined above shows no consistent positive value for helping students generally, or particular groups of students, to learn better. Taking all studies into account, the balance of findings is chiefly of no strong effect either favorable or unfavorable. Among the studies showing significant effects, the slight preponderance of evidence showing the practice favorable for the learning of high ability students is more than offset by evidence of unfavorable effects on the learning of average and low ability groups, particularly the latter. There is no appreciable difference in the effects at elementary and secondary school levels. Finally, those instances of special benefit under ability grouping have generally involved substantial modification of materials and methods, which may well be the influential factors wholly apart from grouping.

The findings regarding impact of ability grouping on the affective development of children are essentially unfavorable. Whatever the practice does to build (inflate?) the egos of children in the high groups is overbalanced by evidence of unfavorable effects of stigmatizing average and low groups as inferior and incapable of learning.

In the absence of evidence of positive effects on learning and personal development of children, and in the light of negative effects on the scholastic achievement and self concepts of low ability groups, the tendency of ability grouping to separate children along ethnic and socioeconomic lines must be deemed to discriminate against children from low socioeconomic classes and minority groups. The mechanism may be said to operate primarily by denying the low groups the scholastic stimulation of their more able peers, and by stigmatizing the low groups as inferior and incapable of learning in their own eyes and those of their teachers. McPartland's data are particularly significant in showing that whatever superior achievement is shown by blacks in desegregated schools, is produced by the superior achievement of blacks in predominantly white (middle class) classroom groups.

Throughout this document we have moved back and forth between ethnic and socioeconomic variables. The fundamental fact of the situation is that minority group membership is consistently and strongly associated with low socioeconomic status. Conversely, high socioeconomic status is strongly associated with membership in the predominant "white" culture. It has not seemed practical or profitable to attempt to delineate these effects differentially. The practical

circumstance is that minority groups preponderantly suffer the disadvantages of low socioeconomic status, increased by the fact of being more immediately identifiable by physical appearance. One can only hope that continuing attention will be given to the socioeconomic factor as basic.

Four brief footnotes. First, ability grouping is undesirable even where ethnic and socioeconomic factors are present, as they generally are. Second, removal of ability grouping has no effect on ethnic discrimination where population movement has already produced ethnic isolation. Third, studies of other minority groups than blacks are needed to bring proper attention to the plight of these smaller minority groups, whose present situation is quite as serious, but not as prominent. Fourth, socioeconomic isolation needs to be elevated to central attention.

Finally, nothing included here may be taken as conclusive evidence that a plan of classroom organization and related procedures may not be effective if well designed to achieve its purpose--for gifted, for mentally retarded, or children generally. The evidence simply indicates that ability grouping *per se* tends to be ineffective and do more harm than good. Any procedure that involves ability grouping and corollary ethnic separation must be justified in terms of other strong evidence of likely beneficial effects.

References

Adkison, Marion Robert. "A Comparative Study of Pupil Attitudes under Conditions of Ability and of Heterogeneous Grouping." Claremont, California: Claremont Graduate School, 1964. Abstract: Dissertation Abstracts 28:3869-A, April 1968.

Antonovsky, Aaron and Lerner, M. J. "Negro and White School Youth in Elmira." In Discrimination and Low Income, Aaron Antonovsky and Lewis L. Lorwin, eds. New York: New School for Social Research, 1959. pp. 103-45.

Bacher, Jesse Hiram. "The Effect of Spcial Class Placement on the Self-Concept, Social Adjustment, and Reading Growth of Slower Learners." New York: New York University, 1964. Abstract: Dissertation Abstracts 25:7071, June 1965.

Barber, Theodore X., Calverley, David S., Forgione, Albert, McPeake, John D., Chaves, John F., and Bowen, Barbara. Five Attempts to Replicate the Experimenter Bias Effect. Journal of Consulting and Clinical Psychology 33(1):1-6, February 1969.

Barker Lunn, Joan C. Streaming in the Primary School. London: National Foundation for Educational Research in England and Wales, 1970.

Bernstein, S. and Esposito, D. "On Grouping in the Experimental Elementary School Project," paper prepared for ERIC Information Retrieval Center for the Disadvantaged, Teachers College, Columbia University, November 1969.

Billett, R. O. The Administration and Supervision of Homogeneous Groupings. Columbus, Ohio: Ohio State University, 1932.

Borg, Walter R. Ability Grouping in the Public Schools, Second Edition Madison, Wisconsin: Dembar Educational Research Services, Inc., 1966.

Borg, Walter R. and Pepich, Tony. "Grouping of Slow Learning High School Pupils." Journal of Secondary Education, 41:231-38; May 1966.

Borg, Walter R. and Maxfield, M. Richard. "A Study of the Effects of Different Kinds of Ability Grouping on Personal Relationships Among High School Students." Washington, D.C.: U. S. Office of Education, Bureau of Research, Project 5-8042 December 1967. Abstract: Research in Education ED 017 002, 1968.

Bouri, Janet and Barker Lunn, Joan. Too Small to Stream. A Study of Grouping in Small Junior High Schools. Occasional Publications Series No. 21. London: National Foundation for Educational Research in England and Wales, 1969.

Brophy, Jere E. and Good, T. L. "Teachers' Communication of Differential Expectations for Children's Classroom Performance: Some Behavioral Data" Research and Development Center in Teacher Education, University of Texas at Austin, Report Series No. 25, 1969 (Also, in press, Journal of Educational Psychology 1970)

Byers, Loretta. "Ability Grouping - Help or Hindrance to Social and Emotional Growth." School Review, 69:449-56, Winter 1961.

Carter, Thomas P. Mexican Americans in School: A History of Educational Neglect. New York: College Entrance Examination Board, 1970.

Clark, Kenneth B. "Educational Stimulation of Racially Disadvantaged Children." In Education in Depressed Areas; A. Harry Passow, ed. New York: Bureau of Publications, Teachers College, Columbia University, 1963.

Coleman, Hubert A. "The Relationship of Socio-economic Status to the Performance of Junior High School Students." Journal of Experimental Education, 9:61-63; September 1940.

Coleman, James S., Campbell, Ernest Q., Hobson, Carol J., McPartland James, Mood, Alexander M., Weinfeld, Frederic D., and York, Robert L. Equality of Educational Opportunity. Washington, D.C.: Government Printing Office, 1966.

Cox, F. N. "Educational Streaming and General and Test Anxiety." Child Development, 33:381-90, June 1962.

Daniels, J. C. "The Effects of Streaming in the Primary School. II. A Comparison of Streamed and Unstreamed Schools." British Journal of Psychology, 31:119-27, June 1961.

De Grow, Gerald S. "A Study of the Effects of the Use of Vertical Reading Ability Grouping for Reading Classes as Compared with Heterogeneous Groupings in Grades Four, Five, and Six in the Port Huron Area Public Schools of Michigan over a Three-Year Period." Ann Arbor, Michigan: University of Michigan, 1963. Abstract: Dissertation Abstracts 24:3166, February 1964.

Deitrich, Francis R. "Comparison of Sociometric Patterns of Sixth-Grade Pupils in Two School Systems: Ability Grouping Compared with Heterogeneous Grouping." Journal of Educational Research, 57:507-13, July 1964.

Deutsch, Martin and Brown, Bert. "Social Influences in Negro-White Intelligence Differences." The Journal of Social Issues, 20:24-36, April 1964.

Douglas, J. W. B. The Home and the School: A Study of Ability and Attainment in the Primary School. London: MacGibbon and Kee, 1964.

Dreger, R. M. and Miller, K. S. "Comparative Psychological Studies of Negroes and Whites in the United States." Psychological Bulletin. 57:361-402, September 1960.

Drews, E. M. Student Abilities, Grouping Patterns and Classroom Interactions. East Lansing, Michigan: Office of Research and Publications, Michigan State University, 1963.

Dyson, Ernest. "A Study of the Relationships Between Acceptance of Self, Academic Self-Concept, and Two Types of Grouping Procedures Used with Seventh Grade Pupils." Philadelphia: Temple University, 1965. Abstract: Dissertation Abstracts 26:1475-76, September 1965.

Eash, Maurice J. "Grouping: What Have We Learned?" Educational Leadership, 18:429-34; April 1961.

Ekstrom, Ruth B. Experimental Studies of Homogeneous Grouping: A Review of the Literature. Princeton, New Jersey: Educational Testing Service, 1959.

Fick, Wayne W. "The Effectiveness of Ability Grouping in Seventh Grade Core Classes." Lawrence, Kansas: University of Kansas, 1962. Abstract: Dissertation Abstracts 23:2753, February 1963.

Flowers, Charles Edward. "Effects of an Arbitrary Accelerated Group Placement on the Tested Academic Achievement of Educationally Disadvantaged Students." New York: Columbia University, 1966. Abstract: Dissertation Abstracts 27:991-A, October 1966.

Franseth, Jane. A Guide to Research and Informed Judgment on Grouping Children, Education Briefs No. 40. Washington, D.C.: Office of Education, U. S. Department of Health, Education, and Welfare, May 1964.

Goldberg, M. L. "Factors Affecting Educational Attainment in Depressed Areas." In Education in Depressed Areas, A. Harry Passow, ed. New York: Bureau of Publications, Teachers College, Columbia University, 1963. pp. 68-99.

Goldberg, Miriam L., Passow, A. Harry, and Justman, Joseph. The Effects of Ability Grouping. New York: Teachers College Press, Columbia University, 1966.

Goldstein, Bernard. Low Income Youth in Urban Areas. A Critical Review of the Literature. New York: Holt, Rinehart, and Winston, Inc., 1967.

Goldworth, Mary. "The Effects of an Elementary School Fast-Learner Program on Children's Social Relationships." Exceptional Children, 26:59-63, October 1959.

Good, T. L. and Brophy, J. E. "Do Boys and Girls Receive Equal Opportunity in the First Grade Reading Instruction?" Research and Development Center in Teacher Education, University of Texas at Austin, Report Series No. 24, 1969.

Goodlad, John I. "Classroom Organization." In Encyclopedia of Educational Research, Third Edition, Chester W. Harris, ed. New York: The Macmillan Company, 1960. pp. 221-25.

Griffin, A. "Effects of Secondary School Organization on the Development of Intellectual Attainmentsin England and Attitudes to School." British Journal of Educational Psychology, 39:191, June 1969.

Havighurst, Robert J. Information from the Study of American Indian Education contained in a letter to the project committee, dated February 13, 1970.

Heathers, Glen. "Grouping." In Encyclopedia of Educational Research, Fourth Edition, Robert L. Ebel, ed. New York: The Macmillan Company, 1969. pp. 559-70.

Hobson v. Hansen. Congressional Record, June 21, 1967.

Justman, Joseph. A Comparison of the Functioning of Intellectually Gifted Children Enrolled in Special Progress and Normal Progress Classes in Junior High School. Doctoral thesis. Columbia University, 1953.

Kariger, Roger H. "The Relationship of Lane Grouping to the Socioeconomic Status of the Parents of Seventh-Grade Pupils in Three Junior High Schools." East Lansing, Michigan: Michigan State University, 1962. Abstract: Dissertation Abstracts 23:4586, June 1963.

Kellmer-Pringle, M. L. and Cox, T. "The Influence of Schooling and Sex on Test and General Anxiety as Measured by Sarason's Scales." Journal of Child Psychology, Psychiatry. 4:157-65, December 1963.

Kennedy, Wallace A., Van De Riet, Vernon, and White, James C., Jr. "A Normative Sample of Intelligence and Achievement of Negro Elementary School Children in the Southeastern United States." Monographs of the Society for Research in Child Development 28:6, 1963.

Kline, Rosemary Elizabeth. "A Longitudinal Study of the Effectiveness of the Track Plan in the Secondary Schools of a Metropolitan Community." St. Louis: St. Louis University, 1963. Abstract: Dissertation Abstracts 25:324, July 1964.

Levy, Philip, Gooch, Stan, and Kellmer-Pringle, M. L. "A Longitudinal Study of the Relationship between Anxiety and Streaming in a Progressive and a Traditional Junior School." British Journal of Educational Psychology. 39:166-73, June 1969.

Locke, Carolyn Jane. "The Effect of Two Plans of Elementary School Organization on Rapid Learners and Non-Rapid Learners." Denton, Texas: North Texas State University, 1962. Abstract: Dissertation Abstracts 23:1289, October 1962.

Loomer, Bradley Max. "Ability Grouping and Its Effect Upon Individual Achievement." Iowa City, Iowa: State University of Iowa, 1962. Abstract: Dissertation Abstracts 23:1581, November 1962.

Luchins, Abraham S. and Luchins, Edith H. "Children's Attitudes Toward Homogeneous Groupings." Pedogogical Seminary and Journal of Genetic Psychology, 72:3-9, March 1948.

Lutrell, Jack. "A Comparative Investigation of the Academic Achievement and Personality Development of Gifted Sixth-Grade Pupils in a Special Class and in Regular Classrooms in the Public Schools of Greensboro, North Carolina." Chapel Hill, North Carolina: University of North Carolina, 1958. Abstract: Dissertation Abstracts 19:2536, April-June 1959.

Mann, Horace. "How Real Are Friendships of Gifted and Typical Children in a Program of Partial Segregation?" Exceptional Children, 23:199-201, 206; February 1957.

Mann, Maxine. "What Does Ability Grouping Do to the Self-Concept?" Childhood Education, 36:356-60; April 1960.

Matzen, Stanley P. The Relationship between Racial Composition and Scholastic Achievement in Elementary School Classrooms. Palo Alto: Stanford University, 1965. Abstract: Dissertation Abstracts 26:6475, May 1966.

Mayeske, George, Wisler, Carol E., Beaton, Albert E. Jr., Weinfeld, Frederic, Cohern, Wallace M., Okoda, Tetsuo, Proshek, John M., and Tabler, Kenneth S. A Study of Our Nation's Schools. Washington, D.C.: Government Printing Office, 1969.

Mayeske, George. Teacher Attributes and School Achievement. Paper presented at a U. S. Office of Education Bureau of Education Development Symposium, February 12, 1970.

Maynor, Waltz. "Academic Performance and School Integration: A Multi-Ethnic Analysis" Unpublished dissertation, Duke University, 1970.

McCandless, B. R. Children: Behavior and Development. New York: Holt, Rinehart and Winston, Inc., 1967.

McDill, Mary S., Stinchcombe, Arthur L., and Walker, Dollie. "Segregation and Educational Disadvantage: Estimates of the Influence of Different Segregating Factors." Sociology of Education 41:239-46, Summer 1968.

McPartland, James. "The Relative Influence of School Desegregation and of Classroom Desegregation on the Academic Achievement of Ninth Grade Negro Students." Journal of Social Issues, 25(3):93-102, Summer 1969.

McPartland, James. The Segregated Student in Desegregated Schools. Sources of Influence on Negro Secondary Students. Baltimore, Maryland: John Hopkins University, June 1968.

Mehl, Robert Franklin, Jr. "A Study of Relationships between Homogeneous Grouping in the School and the Social Class Structure in an Up-State New York Community." Albany, New York: State University of New York at Albany, 1965. Abstract: Dissertation Abstracts 27:4085-A, June 1967.

Mikkelson, James Eliot. "An Experimental Study of Selective Grouping and Acceleration in Junior High School Mathematics." St. Paul, Minnesota: University of Minnesota, 1962. Abstract: Dissertation Abstracts 23:4226, May 1963.

Miner, B. "Sociological Background Variables Affecting School Achievement." Journal of Educational Research. 61:372-81, April 1968.

Morgenstern, Anne. "A Comparison of the Effects of Heterogeneous and Homogeneous (Ability) Grouping on the Academic Achievement and Personal-Social Adjustment of Sixth-Grade Children." New York: New York University, 1963. Abstract: Psychological Abstracts 24:1054, September 1963.

National Education Association, Research Division. Ability Grouping, Research Summary 1968-S3. Washington, D.C.: National Education Association, 1968.

New York State Education Department, Division of Research. Racial and Social Isolation in the Schools (Preliminary Report). Albany, New York: Office of Public Information, New York State Education Department, 1969.

Ogletree, Earl. "Homogeneous Ability Grouping-British Style. Peabody Journal of Education, 47:20-25, July 1969.

Olivarri, Martin C. "Some Relationships of Ability Grouping to Student Self-Concept." Berkeley, University of California, 1967. Abstract: Dissertation Abstracts 28:2518A, January 1968.

Otto, Henry. "Classification of Pupils." In Encyclopedia of Educational Research, Second Edition, Walter S. Monroe, ed. New York: The Macmillan Company, 1950. pp. 376-78.

Otto, Henry J. "Elementary Education - II, Organization and Administration." In Encyclopedia of Educational Research, First Edition, Walter S. Monroe, ed. New York: The Macmillan Company, 1941. pp. 428-46.

Passow, A. Harry. Toward Creating a Model Urban School System: A Study of the Washington, D.C. Public Schools. New York: Teachers College, Columbia University, 1967.

Peterson, Richard Lee. "An Experimental Study of the Effects of Ability Grouping in Grades Seven and Eight." St. Paul: University of Minnesota, 1966. Abstract: Dissertation Abstracts 28:103-A, July 1967.

Plainfield Study. Grouping Students for Instruction in the Plainfield, New Jersey, School System. New York: Institute of Field Studies, Teachers College, Columbia University, July 1969.

Provus, Malcolm M. "Ability Grouping in Arithmetic." Elementary School Journal, 60:391-98, April 1960.

Rosenthal, Robert. Empirical vs. Decreed Validation of Clocks and Tests. American Educational Research Journal 6:689-691, November 1969.

Rosenthal, Robert, and Jacobson, Lenore. Pygmalion in the Classroom. New York: Holt, Rinehart and Winston, Inc., 1968.

Rudd, W. G. "The Psychological Effect of Streaming by Attainment." British Journal of Educational Psychology, 28:47-60, February 1958.

Sarthory, Joseph A. "The Effects of Ability Grouping in the Multi-Cultural School Situation." Albuquerque: University of New Mexico, 1968. Abstract: Dissertation Abstracts 29:451-A, August 1968.

Sexton, Patricia Cayo. Education and Income; inequalities of opportunity in our public schools. New York: The Viking Press, Inc., 1961

Simpson, R. E. and Martinson, Ruth A. Educational Programs for Gifted Pupils. Sacramento, California: California State Department of Education, January 1961.

Stinchcombe, Arthur L., McDill, Mary S., and Walker, Dollie R. "Demography of Organizations" American Journal of Sociology, 74:221-29, November 1968.

Stinchcombe, Arthur L., McDill, Mary S., and Walker, Dollie R. "Is There a Racial Tipping Point in Changing Schools?" Journal of Social Issues 25(1):127-36, Winter 1969.

Thorndike, Robert L. But You Have to Know How to Tell Time. American Educational Research Journal 6:692, November 1969.

Thorndike, Robert L. Review of "Pygmalion in the Classroom." American Educational Research Journal 5:708-711, November 1968.

Tobin, John Francis. "An Eight-Year Study of Classes Grouped Within Grade Levels on the Basis of Reading Ability." Boston: Boston University, 1965. Abstract: Dissertation Abstracts 26:5141, March 1966.

Torgelson, John Winston. A Comparison of Homogeneous and Heterogeneous Grouping for Below-Average Junior High School Students. Minneapolis: University of Minnesota, 1963. Abstract: Dissertation Abstracts 25:2300-01, October 1964.

U. S. Commission on Civil Rights. Racial Isolation in the Public Schools, I and II. Washington, D.C.: U. S. Government Printing Office, 1967.

U. S. National Advisory Commission on Civil Disorders. Report of the National Advisory Committee on Civil Disorders. Washington, D.C.: Government Printing Office, 1968.

Walker, Dollie R., Stinchcombe, Arthur L., and McDill, Mary S. School Desegregation in Baltimore. Baltimore, Maryland: Center for the Study of Social Organization of Schools, Johns Hopkins University, August 1967.

Weinberg, Meyer. Desegregation Approach: An Appraisal. Second Edition, Bloomington, Indiana: Phi Delta Kappa, May 1970.

Wilcox, John. "A Search for the Multiple Effects of Grouping Upon the Growth and Behavior of Junior High School Pupils." Ithaca, New York: Cornell University, 1963. Abstracts: Dissertation Abstracts 24:205, July 1963.

Wilhelms, Fred T. and Westby-Gibson, Dorothy. "Grouping: Research Offers Leads." Educational Leadership, 18:410-13, 476, April 1961.

Willcutt, Robert Ernest. Ability Grouping by Content Subject Areas in Junior High School Mathematics. Bloomington: Indiana University, 1967. Abstract: Dissertation Abstracts 28:2152A, December 1967.

Willig, C. J. "Social Implications of Streaming in the Junior School." Educational Research, 5:151-54; February 1963.

Wilson, Alan B. "Educational Consequences of Segregation in a California Community." In Racial Isolation in the Public Schools II, A Report of the U. S. Commission on Civil Rights. Washington, D.C.: Government Printing Office, 1967. pp. 165-206.

Zweibelson, I., Bahmuller, M., and Lyman, L. "Team Teaching and Flexible Grouping in the Junior High School Social Studies." The Journal of Experimental Education, 34:20-32; Fall 1965.